SEO 实战密码——60 天网站流量提高 20 倍（第 2 版）

昝辉 Zac　著

電子工業出版社
Publishing House of Electronics Industry
北京・BEIJING

内 容 简 介

本书是畅销书升级版，详细和系统地介绍了正规、有效的 SEO 实战技术，包括为什么要做 SEO、搜索引擎工作原理、关键词研究、网站架构优化、外部链接建设、SEO 效果检测及策略修正，SEO 作弊及惩罚、排名因素列表、常用的 SEO 工具、SEO 项目管理中需要注意的问题等专题，最后提供了一个非常详细的案例供读者参考。第 2 版对第 1 版中不再适用的内容进行了小幅调整，如删除已无法使用的工具，更新已产生变化的搜索结果页面抓图等，并且增加了 2011 年 Google 排名因素调查的最新结果等内容。

本书不仅对需要做 SEO 的人员有帮助，如个人站长、公司 SEO 或网络营销人员、SEO 服务公司人员等，对所有从事与网站相关工作的人都有参考价值，如网站设计人员、程序员、大中专院校网络营销和电子商务专业学生、网络公司技术和营销团队、传统商业公司电子商务团队、新媒体营销从业人员等。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

SEO 实战密码：60 天网站流量提高 20 倍 / 昝辉著. —2 版. —北京：电子工业出版社，2012.5
ISBN 978-7-121-16767-6

Ⅰ. ①S… Ⅱ. ①昝… Ⅲ. ①互联网络－情报检索②电子商务－网站－基本知识 Ⅳ. ①G354.4 ②F713.36

中国版本图书馆 CIP 数据核字（2012）第 070083 号

责任编辑：李　冰
印　　刷：北京天宇星印刷厂
装　　订：三河市皇庄路通装订厂
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　邮编 100036
开　　本：787×1092　1/16　印张：36.5　字数：908 千字　彩插：4
印　　次：2012 年 5 月第 1 次印刷
印　　数：4 000 册　定价：89.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

PREFACE

I enjoyed doing an interview with Zac about search engine optimization (SEO) back in 2007. Not only did Zac ask great questions, but he has provided countless people with helpful, solid advice over the years. So when Zac asked me to write a forward for his book about SEO, I was happy to say yes.

I think learning about SEO can be good for anyone who works with the web. Not only designers and programmers, but also CEOs and regular users can benefit from knowing more about how search engines rank pages, and why some pages rank more highly than others. SEO can be done in a good way that keeps users' needs in mind and create useful websites in alignment with search engine quality guidelines. So SEO can be a powerful tool that not only helps a website rank higher but also makes a website easier to use.

Some people think that SEO only means spam or deceptive techniques, and that's not true. SEO can include designing a website or web page to be clear and easy for people and computers to discover new pages by following links. Attention to SEO can suggest phrases that people will type when looking for your products or pages, which you can then include on the page in a natural way. Learning SEO includes learning the lesson that people want to read high-quality information and that they appreciate useful services or resources. Students of SEO also learn ways to promote their site in a number of ways that can raise awareness and result in more links to a web page.

It turns out that the Chinese web is different from the English web or the German web. Different countries have different link structures, not to mention different keyword areas that are more or less popular. Countries also have different mixes between standalone domain names vs. content that appears on forums or bulletin boards. For that reason, it's helpful to have an SEO book that is written specifically for the Chinese market. I'm glad that Zac has written that book.

Matt Cutts, head of webspam team for Google

序

早在 2007 年，我愉快地与 Zac 进行过一次关于搜索引擎优化（SEO）的访谈。Zac 不仅问了很好的问题，而且也在这些年为无数人提供了有益、扎实的建议。所以，当 Zac 请我为他的《SEO 实战密码——60 天网站流量提高 20 倍》一书写序时，我很高兴地说“Yes”。

我认为学习 SEO 对任何从事网络工作的人都是好事。不仅是设计师和程序员，CEO 和普通用户如果更多地了解搜索引擎怎样排名、为什么有的网页比其他的排名更高，都能受益良多。

SEO 能以很合理的方式进行，既照顾到用户需求，又创造出有用的、符合搜索引擎质量指南的网站。SEO 是个强有力的工具，既能帮助网站提高排名，又使网站容易使用。

有的人认为 SEO 只意味着发垃圾和欺骗性的手段，这是不正确的。SEO 可以包括为用户设计一个清晰易用的网站，电脑可以通过跟踪链接发现新的页面。关注 SEO 可以发现用户寻找产品或网页时输入的关键词，然后你可以在页面上自然融入这些词。学习 SEO 使你了解人们需要高质量的信息，并且喜欢有用的服务和资源。学 SEO 的人也能学习到提高认知度并且为网页带来更多链接的各种网站推广方法。

事实证明，中文网络与英文或德文网络不同。不同国家的网站有不同的链接结构，更不要说不同的流行关键词。不同的国家也有不同的独立域名和出现在论坛或电子公告板的内容组合。因此，有一本专门为中国市场写的 SEO 书是很有帮助的。很高兴 Zac 写了这样一本书。

Matt Cutts，Google 反垃圾组负责人

第 2 版前言

几年前和一位擅长策划出版 IT 类书籍的朋友聊天，他告诉我，现在的 IT 类书能卖一万本就算畅销书了。所以，我的两本密码书（《SEO 实战密码》第 1 版于 2010 年 12 月上市、《网络营销实战密码》于 2009 年 2 月上市）能卖这么好，并且能获得行业内的畅销图书奖，还挺出乎我意料的。现在有些公司在招聘资深 SEO 分析师时，甚至要将对《SEO 实战密码》融会贯通作为招聘条件之一，都让我很意外。

出版社编辑说，两本密码书已经至少服务了 10 万名以上的读者，能够经过如此数量级的读者检验，这在技术书领域已接近奇迹了。有多少读者看了这两本书其实无法确知。一方面，应该有相当多的人两本书都买了，购书总人数会比书的发行量小。另一方面，一本书有多个人看也很正常，我在参观朋友的公司时经常看到我的书在公司桌上放着，供大家一起参考。这还不算无数盗版。

总之，这两本书在同类书中销售算很不错的。

这倒不是因为书写得有多好。我觉得主要有两个原因：一是 SEO 和站长圈子里知道我的人不少，大家都很捧场帮忙。二是市场需求比较大。可以这样说，做网站的，无论是商业公司还是个人站长，没有没听说过 SEO 的。

三四年前，SEO 行业很喧嚣，这两年消停了很多。不过我倒觉得，沉寂了的 SEO 有更大的市场。一个行业，或者一种技术，一方面藏着掖着，故作神秘，或者被神秘，另一方面，每个人都在热烈讨论它，热闹非凡，这往往不是什么好事。三四年前的 SEO 就是这样。

现在，很少看到令人耳目一新的 SEO 博客或论坛帖子了，老 SEO 博客们几乎都不更新了，也没有什么令人瞩目的新 SEO 博客出现，点石的创始人和老版主们大多不再以 SEO 为职业了，认为 SEO 实在没什么技术含量的人越来越多了……

但，SEO 已经成为大部分网站的标配。

这对所有的 SEO 从业人员和爱好者都是好事。SEO 成为一个没人再热烈讨论但每个网站都需要的技能和职位。

特色内容

看过几本 SEO 书及网上很多新手的反馈意见后觉得，一些刚开始学 SEO 的人需要

一个手把手示范的过程。小范围单独指导是个方法，但无法扩展，要对更多的人有益，还得靠书籍。所以本书包含了一个非常详细、篇幅近 6 万字的真实案例。这是本书独特的地方，在其他地方还没有见到过这样的案例。

本书详细和系统地介绍了正规、有效的 SEO 实战技术，包括关键词研究、网站架构优化、页面优化、外链建设、效果监测及策略修正，以及作弊与惩罚、排名因素列表等专题。

第 1 章　为什么要做 SEO	讨论为什么要做 SEO
第 2 章　了解搜索引擎	介绍搜索引擎工作原理，为深入了解 SEO 打下良好基础
第 3 章　竞争研究	讨论竞争研究，包括对关键词、竞争对手及自己网站的深入研究
第 4 章　网站结构优化 第 5 章　页面优化	介绍站内优化，包括网站结构和页面优化
第 6 章　外部链接建设	探讨外部链接建设
第 7 章　SEO 效果监测及策略修改	介绍 SEO 效果监测及策略调整
第 8 章　SEO 作弊及惩罚	介绍 SEO 作弊及搜索引擎惩罚
第 9 章　SEO 专题 第 10 章　SEO 观念及原则	讨论不好归类的一些专题，包括 SEO 观念、垂直搜索的排名等
第 11 章　SEO 工具	介绍常用的 SEO 工具
第 12 章　SEO 项目管理	简单讨论 SEO 项目管理中需要注意的问题
第 13 章　搜索引擎排名因素调查	Google 和百度排名因素调查及列表，供 SEO 人员快速参考
第 14 章　SEO 案例分析	真实案例
附录 A　SEO 术语	SEO 相关术语

读者对象

我相信这本书不仅对需要做 SEO 的人有帮助，如个人站长、公司 SEO 或网络营销人员、SEO 服务公司人员，对所有从事与网站相关工作的人都有参考价值，如网站设计人员、程序员、大专院校网络营销和电子商务专业学生、网络公司技术和营销团队、传统商业公司电子商务团队等，因为 SEO 已经是所有网站的基本要求。

致谢

这本书引领了并将继续引领更多的朋友走进 SEO，我衷心地高兴。感谢读者和朋友们这些年来各种形式的支持、鼓励和帮助。

Zac

2012 年 4 月

CONTENTS

目　录

第 1 章　为什么要做 SEO ····· 1

1.1　什么是 SEO ····· 1

1.2　为什么要做 SEO ····· 2

1.3　搜索引擎简史 ····· 6

第 2 章　了解搜索引擎 ····· 14

2.1　搜索引擎与目录 ····· 15

2.2　搜索引擎面对的挑战 ····· 15

2.3　搜索结果显示格式 ····· 17

2.3.1　搜索结果页面 ····· 17

2.3.2　经典搜索结果列表 ····· 20

2.3.3　整合搜索结果 ····· 21

2.3.4　缩进列表 ····· 22

2.3.5　全站链接 ····· 22

2.3.6　迷你全站链接 ····· 23

2.3.7　One-box ····· 23

2.3.8　富摘要 ····· 23

2.3.9　面包屑导航 ····· 24

2.3.10　说明文字中的链接 ····· 24

2.4　搜索引擎工作原理简介 ····· 24

2.4.1　爬行和抓取 ····· 25

2.4.2　预处理 ····· 28

2.4.3　排名 ····· 33

2.5　链接原理 ····· 37

2.5.1　李彦宏超链分析专利 ····· 38

2.5.2　HITS 算法 ····· 39

2.5.3　TrustRank 算法 ····· 40

2.5.4　Google PR ····· 41

2.5.5　Hilltop 算法 ····· 45

2.6　用户怎样浏览和点击搜索结果 ····· 46

2.6.1　英文搜索结果页面 ····· 46

2.6.2　中文搜索结果页面 ····· 50

2.6.3　整合搜索及个人化搜索 ····· 52

2.7　高级搜索指令 ····· 54

2.7.1　双引号 ····· 54

2.7.2　减号 ····· 55

2.7.3　星号 ····· 56

2.7.4　inurl: ····· 57

2.7.5　inanchor: ····· 57

2.7.6　intitle: ····· 58

2.7.7　allintitle: ····· 59

2.7.8　allinurl: ····· 59

2.7.9　filetype: ····· 60

2.7.10　site: ····· 60

2.7.11　link: ····· 61

2.7.12　linkdomain: ····· 62

2.7.13　related: ····· 62

2.7.14　综合使用高级搜索指令 ····· 63

第 3 章　竞争研究 ····· 65

3.1　为什么要研究关键词 ····· 65

3.1.1　确保目标关键词有人搜索 ····· 65

3.1.2　降低优化难度 ····· 65

3.1.3　寻找有效流量 ····· 66

3.1.4　搜索多样性 ····· 66

3.1.5　发现新机会 ····· 67

3.2　关键词的选择 ····· 67

3.2.1　内容相关 ····· 67

3.2.2　搜索次数多，竞争小 ····· 68

3.2.3　主关键词不可太宽泛 ····· 68

3.2.4　主关键词也不可太特殊 ····· 68

3.2.5　商业价值 ····· 69

3.3　关键词竞争程度判断 ····· 69

3.3.1　搜索结果数 ····· 69

3.3.2　intitle 结果数 ····· 70

3.3.3 竞价结果数 ……70
3.3.4 竞价价格 ……71
3.3.5 竞争对手情况 ……72
3.3.6 内页排名数量 ……72
3.4 核心关键词 ……73
3.4.1 头脑风暴 ……73
3.4.2 同事、朋友 ……73
3.4.3 竞争对手 ……73
3.4.4 查询搜索次数 ……74
3.4.5 确定核心关键词 ……75
3.5 关键词扩展 ……77
3.5.1 关键词工具 ……77
3.5.2 搜索建议 ……78
3.5.3 相关搜索 ……78
3.5.4 其他关键词扩展工具 ……79
3.5.5 各种形式的变体 ……79
3.5.6 补充说明文字 ……80
3.5.7 网站流量分析 ……81
3.5.8 单词交叉组合 ……81
3.6 关键词分布 ……81
3.6.1 金字塔形结构 ……81
3.6.2 关键词分组 ……82
3.6.3 关键词布局 ……83
3.6.4 关键词-URL 对应表 ……83
3.7 长尾关键词 ……84
3.7.1 长尾理论 ……84
3.7.2 搜索长尾 ……85
3.7.3 怎样做长尾关键词 ……86
3.8 三类关键词 ……86
3.8.1 导航类关键词 ……86
3.8.2 交易类关键词 ……88
3.8.3 信息类关键词 ……88
3.9 预估流量及价值 ……88
3.9.1 确定目标排名 ……89
3.9.2 预估流量 ……89
3.9.3 预估搜索流量价值 ……93
3.10 关键词趋势波动和预测 ……93
3.10.1 长期趋势 ……93
3.10.2 季节性波动 ……94
3.10.3 社会热点预测 ……95
3.11 竞争对手研究 ……97
3.11.1 域名权重相关数据 ……97
3.11.2 网站优化情况 ……99
3.11.3 网站流量 ……100
3.12 快速网站诊断 ……100
3.12.1 robots 文件检查 ……101
3.12.2 首选域设置 ……102
3.12.3 关键词排名 ……103
3.12.4 外部链接 ……104
3.12.5 网站内容 ……105
3.12.6 内部链接 ……106
3.12.7 抓取错误及统计 ……107
3.12.8 HTML 建议 ……108
3.12.9 模拟蜘蛛抓取 ……109
3.12.10 网站性能 ……110
第 4 章 网站结构优化 ……112
4.1 搜索引擎友好的网站设计 ……113
4.2 避免蜘蛛陷阱 ……118
4.2.1 Flash ……118
4.2.2 Session ID ……119
4.2.3 各种跳转 ……119
4.2.4 框架结构 ……119
4.2.5 动态 URL ……120
4.2.6 JavaScript 链接 ……120
4.2.7 要求登录 ……120
4.2.8 强制使用 Cookies ……121
4.3 物理及链接结构 ……121
4.3.1 物理结构 ……121
4.3.2 链接结构 ……122
4.4 清晰导航 ……123
4.5 子域名和目录 ……125
4.6 禁止收录机制 ……126
4.6.1 robots 文件 ……126
4.6.2 meta robots 标签 ……129
4.7 nofollow 的使用 ……130
4.8 URL 静态化 ……131
4.8.1 为什么静态化 ……131
4.8.2 怎样静态化 URL ……133
4.8.3 URL 不需要静态化吗 ……133

4.9 URL 设计 ····· 135
4.10 网址规范化 ····· 137
4.10.1 为什么出现不规范网址 ··· 137
4.10.2 网址规范化问题 ····· 138
4.10.3 解决网址规范化问题 ····· 138
4.10.4 301 转向 ····· 139
4.10.5 Canonical 标签 ····· 141
4.11 复制内容 ····· 142
4.11.1 产生复制内容的原因 ····· 142
4.11.2 复制内容的害处 ····· 144
4.11.3 消除复制内容 ····· 144
4.12 绝对路径和相对路径 ····· 145
4.12.1 绝对路径 ····· 146
4.12.2 相对路径 ····· 147
4.13 网站地图 ····· 147
4.13.1 HTML 网站地图 ····· 148
4.13.2 XML 网站地图 ····· 148
4.14 内部链接及权重分配 ····· 150
4.14.1 重点内页 ····· 150
4.14.2 非必要页面 ····· 150
4.14.3 大二级分类 ····· 151
4.14.4 翻页过多 ····· 152
4.14.5 单一入口还是多入口 ····· 153
4.14.6 相关产品链接 ····· 154
4.14.7 锚文字分布及变化 ····· 155
4.14.8 首页链接 NoFollow ····· 155
4.14.9 深层链接 ····· 156
4.14.10 分类隔离 ····· 156
4.15 CMS 系统 ····· 157
4.16 404 页面 ····· 159
4.16.1 404 错误代码 ····· 160
4.16.2 404 页面设计 ····· 160
4.16.3 404 错误与外链 ····· 161

第 5 章 页面优化 ····· 162

5.1 页面标题 ····· 162
5.1.1 独特不重复 ····· 162
5.1.2 准确相关 ····· 165
5.1.3 字数限制 ····· 165
5.1.4 简练通顺，不要堆砌 ····· 166
5.1.5 关键词出现在最前面 ····· 167
5.1.6 吸引点击 ····· 167
5.1.7 组合两三个关键词 ····· 168
5.1.8 公司或品牌名称 ····· 169
5.1.9 连词符使用 ····· 169
5.1.10 不要用没有意义的句子 ··· 169
5.1.11 noodp 标签 ····· 169
5.2 描述标签 ····· 170
5.3 关键词标签 ····· 171
5.4 正文中的关键词 ····· 172
5.4.1 词频和密度 ····· 172
5.4.2 前 50～100 个词 ····· 172
5.4.3 关键词变化形式 ····· 173
5.4.4 关键词组临近度 ····· 173
5.4.5 词组的拆分出现 ····· 173
5.4.6 语义分析 ····· 173
5.4.7 分类页面说明文字 ····· 174
5.5 H 标签 ····· 175
5.6 ALT 文字 ····· 176
5.7 精简代码 ····· 177
5.8 内部链接及锚文字 ····· 177
5.9 导出链接及锚文字 ····· 178
5.10 W3C 验证 ····· 178
5.11 黑体及斜体 ····· 178
5.12 页面更新 ····· 179
5.13 Google 沙盒效应 ····· 179

第 6 章 外部链接建设 ····· 181

6.1 外部链接意义 ····· 181
6.1.1 相关性及锚文字 ····· 181
6.1.2 权重及信任度 ····· 182
6.1.3 收录 ····· 182
6.2 Google 炸弹 ····· 183
6.3 链接分析技术 ····· 185
6.4 什么样的链接是好链接 ····· 187
6.5 外部链接查询 ····· 190
6.5.1 链接查询指令 ····· 190
6.5.2 工具查询外链 ····· 191
6.5.3 影响排名的链接 ····· 192
6.6 外部链接原则 ····· 193

6.6.1 难度越大，价值越高……193
6.6.2 内容是根本……194
6.6.3 内容相关性……194
6.6.4 链接来源广泛……194
6.6.5 深度链接……195
6.6.6 锚文字分散自然……195
6.6.7 平稳持续增加……195
6.6.8 质量高于数量……195
6.7 网站目录提交……196
6.7.1 提交前的准备……196
6.7.2 寻找网站目录……197
6.7.3 网站提交……198
6.8 友情链接……199
6.8.1 友情链接页面……199
6.8.2 软件使用……200
6.8.3 寻找交换链接目标……200
6.8.4 交换链接步骤……201
6.8.5 内页正文链接交换……202
6.8.6 交换链接中的小花招……203
6.9 链接诱饵……205
6.9.1 链接诱饵的制作……205
6.9.2 链接诱饵种类和方法……206
6.9.3 链接诱饵之度……220
6.10 其他常规外链建设方法……221
6.11 非链接形式的链接……231
6.12 竞争对手能否通过垃圾外链陷害你……232
6.13 链接工作表……234

第 7 章 SEO 效果监测及策略修改……236

7.1 为什么要监测……236
7.1.1 检验工作成效……236
7.1.2 发现问题，修改策略……236
7.1.3 SEO 完整过程……237
7.2 网站目标设定及测量……237
7.2.1 网站目标……237
7.2.2 网站目标实例……238
7.2.3 网站目标确定原则……239
7.2.4 网站目标影响 SEO 策略……240
7.3 非流量数据监测……240
7.3.1 收录数据……240
7.3.2 排名监测……244
7.3.3 外部链接数据……245
7.3.4 转化和销售……245
7.4 流量数据监测……246
7.4.1 怎样读日志文件……246
7.4.2 常用流量分析工具……249
7.4.3 流量统计分析基础……250
7.5 策略改进……257
7.5.1 收录是否充分……257
7.5.2 哪些页面带来搜索流量……259
7.5.3 目标 URL 排名如何……260
7.5.4 挖掘关键词……261
7.5.5 其他搜索引擎流量……261
7.5.6 长尾效果……263
7.5.7 关键词排名下降……263
7.5.8 链接诱饵成效……264
7.5.9 发现链接伙伴……264
7.5.10 寻找有潜力关键词……264

第 8 章 SEO 作弊及惩罚……266

8.1 白帽、黑帽、灰帽……266
8.1.1 白帽、黑帽是风险度判断……266
8.1.2 道德及法律底线……267
8.1.3 SEO 服务商的底线……267
8.1.4 黑帽 SEO 的贡献……268
8.1.5 承担风险，不要抱怨……268
8.1.6 了解黑帽，做好白帽……269
8.2 主要 SEO 作弊方法……270
8.2.1 隐藏文字（Hidden Text）……270
8.2.2 隐藏链接（Hidden Links）……271
8.2.3 垃圾链接（Link Spam）……271
8.2.4 买卖链接（Paid Links）……273
8.2.5 链接农场（Link Farm）……274
8.2.6 链接向坏邻居（Bad Neighborhood）……275
8.2.7 隐藏页面（Cloaking, Cloaked Page）……275
8.2.8 PR 劫持（PR Hijacking）……276

8.2.9 桥页（Doorway Pages, Bridge Pages）……277
8.2.10 跳转……277
8.2.11 诱饵替换（Bait and Switch）……278
8.2.12 关键词堆积（Keyword Stuffing）……278
8.2.13 大规模站群……279
8.2.14 利用高权重网站……279
8.3 搜索引擎惩罚……281
8.3.1 作弊的积分制……281
8.3.2 不要学大网站……283
8.3.3 不要存侥幸心理……284
8.3.4 搜索引擎惩罚的种类……284
8.3.5 搜索引擎惩罚的检测……285
8.4 被惩罚了怎么办……286
第 9 章 SEO 专题……290
9.1 整合搜索优化……290
9.1.1 什么是整合搜索……290
9.1.2 机会和挑战……291
9.1.3 新闻搜索……292
9.1.4 图片搜索……293
9.1.5 视频搜索……293
9.1.6 地图搜索……294
9.2 更改域名……295
9.3 多个域名的处理……296
9.4 更换服务器……297
9.5 用户行为影响排名……298
9.5.1 用户行为信息收集……298
9.5.2 影响排名的用户行为……298
9.5.3 回归用户体验……300
9.6 域名与 SEO……300
9.6.1 域名后缀……300
9.6.2 域名年龄……301
9.6.3 域名第一次被收录时间……301
9.6.4 域名续费时间……301
9.6.5 域名包含关键词……301
9.6.6 连词符使用……302
9.6.7 品牌优先……302
9.6.8 域名长短……302
9.6.9 域名买卖历史……303
9.6.10 匿名注册信息……303
9.6.11 域名权重……303
9.7 主机与 SEO……303
9.7.1 IP 及整个服务器惩罚……303
9.7.2 服务器设置……304
9.7.3 稳定性……304
9.7.4 主机速度……304
9.7.5 URL 重写支持……304
9.8 多语种内容……305
9.8.1 多语种页面处理……305
9.8.2 当地语言习惯与 SEO……306
9.9 地理定位……306
9.9.1 什么是地理定位……307
9.9.2 地理定位的表现形式……307
9.9.3 地理定位的影响因素……308
9.10 社会化媒体的影响……309
9.10.1 带来链接……309
9.10.2 互动及口碑传播……310
9.10.3 新形式的链接流动成为排名信号……310
9.10.4 网络名誉管理……311
9.11 避免过度优化……311
9.12 SEO 与品牌……312
9.12.1 排名第一就是品牌……312
9.12.2 传统品牌建设与 SEO 结合……313
9.12.3 用 SEO 进行网上危机公关……313
9.13 针对不同搜索引擎的优化……315
9.13.1 SEO 原则不变……315
9.13.2 百度和 Google 的区别……315
9.13.3 英文网站优化……316
9.14 网站改版……317
9.14.1 设计还是 CMS 系统改变……317
9.14.2 不要改 URL……317
9.14.3 分步更改……317
9.15 Google Dance……318
9.15.1 什么是 Google Dance……318

9.15.2 Google 已不再 Dance……319
9.15.3 近年 Google 更新……319
9.16 Google 全站链接……320
9.16.1 全站链接的出现……320
9.16.2 屏蔽全站链接……321
9.16.3 迷你全站链接……321
9.17 个人化搜索……322
9.17.1 什么是个人化搜索……322
9.17.2 个人化搜索对 SEO 的影响……323

第 10 章 SEO 观念及原则……324
10.1 搜索引擎的目标……324
10.1.1 搜索引擎的目标是满足搜索用户……324
10.1.2 搜索引擎不在乎我们……325
10.1.3 搜索引擎在乎垃圾……325
10.2 相关性、权威性、实用性……326
10.2.1 网站内容的相关性……326
10.2.2 网站及网页的权威性……326
10.2.3 网站的实用性……327
10.3 SEO 与赚钱……327
10.3.1 给别人做 SEO……327
10.3.2 给自己做 SEO……329
10.4 SEO 不是免费的……330
10.4.1 人力成本……330
10.4.2 机会成本……331
10.4.3 失败风险……331
10.4.4 SEO 成功风险……331
10.5 不要做奇怪的事……331
10.6 解决基本问题就解决了 95%的问题……332
10.7 自然和平衡的艺术……333
10.8 SEO 是长期策略……334
10.8.1 实施 SEO 需要时间……334
10.8.2 不进则退……335
10.9 没有 SEO 秘籍……335
10.9.1 为什么没有 SEO 秘籍……336
10.9.2 搜索引擎排名算法的秘密……336
10.9.3 SEO 绝招……336
10.10 SEO 不仅是排名……337
10.11 SEO 不是作弊……338
10.12 SEO 与网站运营……339
10.13 内容为王……339
10.13.1 原创内容是 SEO 的根本……339
10.13.2 内容策划是 SEO 策略……340
10.13.3 内容推广……341
10.14 具体问题具体分析……341

第 11 章 SEO 工具……343
11.1 Xenu……344
11.2 Alexa……346
11.3 谷歌趋势……349
11.4 百度指数……352
11.5 百度搜索风云榜……354
11.6 Google Adwords 关键词工具……356
11.7 微软广告工具……358
11.8 Google 百宝箱……360
11.9 Google 快讯……363
11.10 服务器头信息检测器……364
11.11 W3C 验证……365
11.12 雅虎外链检查工具……366
11.13 外链概况工具……368
11.14 IP 地址检查工具……369
11.15 关键词问答……370
11.16 Google 搜索解析……371
11.17 网站索引指数查询……372
11.18 SEO for Firefox……373
11.19 SEO 工具条……375
11.20 火狐浏览器 Search Status 插件……376
11.21 火狐 SeoQuake 插件……378
11.22 站长帮手……381
11.23 关键词排名批量查询工具……382
11.24 SEOmoz 工具……384
11.25 Backlink Watch……387
11.26 TouchGraph……388
11.27 Quintura……388

11.28 Google Ad Planner ····· 389
11.29 Majestic SEO ····· 390
11.30 追词 ····· 397

第 12 章 SEO 项目管理 ····· 400

12.1 内部团队还是 SEO 服务 ····· 400
12.2 寻找 SEO 服务商 ····· 402
12.3 SEO 团队建设 ····· 405
12.4 流程及计划 ····· 407
12.5 绩效考核 ····· 410
12.6 获得高层支持 ····· 411
12.7 沟通、培训及规范 ····· 413
12.8 应急计划 ····· 414

第 13 章 搜索引擎排名因素调查 ····· 417

13.1 Google 排名因素调查 2009 ····· 417
13.1.1 与关键词有关的页面排名因素 ····· 419
13.1.2 与关键词无关的页面排名因素 ····· 420
13.1.3 特定页面链接流行度排名因素 ····· 421
13.1.4 全站链接排名因素 ····· 423
13.1.5 全站非链接排名因素 ····· 423
13.1.6 社会化媒体排名因素 ····· 425
13.1.7 用户数据排名因素 ····· 426
13.1.8 负面排名因素 ····· 426
13.1.9 影响外部链接价值的负面因素 ····· 429
13.1.10 地理定位因素 ····· 430
13.1.11 附加 SEO 数据 ····· 431
13.1.12 链接建设调查 ····· 432
13.2 Google 排名因素调查 2011 ····· 435
13.2.1 域名级别关键词使用 ····· 435
13.2.2 域名级别链接权威度指标 ····· 436
13.2.3 域名级别与关键词无关的指标 ····· 436
13.2.4 域名级别品牌指标 ····· 437
13.2.5 页面级别社会化指标 ····· 437
13.2.6 页面级别链接指标 ····· 438
13.2.7 页面级别关键词使用 ····· 438
13.2.8 页面级别流量数据 ····· 439
13.2.9 页面级别与关键词无关的指标 ····· 440
13.3 百度排名因素调查 2010 ····· 440
13.3.1 与关键词有关的页面排名因素 ····· 441
13.3.2 与关键词无关的页面排名因素 ····· 442
13.3.3 特定页面链接流行度排名因素 ····· 443
13.3.4 全站链接有关排名因素 ····· 444
13.3.5 全站非链接相关排名因素 ····· 445
13.3.6 社会化媒体排名因素 ····· 446
13.3.7 用户数据排名因素 ····· 446
13.3.8 负面排名因素 ····· 447
13.3.9 地理位置定位因素 ····· 448

第 14 章 SEO 案例分析 ····· 452

14.1 竞争对手分析 ····· 453
14.1.1 了解网站基本数据 ····· 453
14.1.2 外部链接 ····· 455
14.1.3 Alexa 数据 ····· 455
14.1.4 Google 趋势流量 ····· 457
14.1.5 网站品牌名称热度 ····· 458
14.1.6 英文比较购物网站情况 ····· 459
14.2 竞争对手网站研究 ····· 460
14.2.1 域名注册信息 ····· 460
14.2.2 基本信息 ····· 460
14.2.3 外部链接 ····· 460
14.2.4 收录 ····· 461
14.2.5 QQ 书签 ····· 462
14.2.6 外链锚文字 ····· 463
14.2.7 网站首页优化 ····· 465
14.2.8 其他页面优化 ····· 470
14.3 亿赐客网站分析 ····· 474
14.3.1 域名注册 ····· 475
14.3.2 Google PR 值 ····· 475

14.3.3 收录……475
14.3.4 外部链接……477
14.3.5 QQ 书签……478
14.3.6 基本流量数据……478
14.3.7 Google 网管工具数据和分析……488
14.4 关键词研究……496
14.4.1 首页……497
14.4.2 分类页面……502
14.4.3 商家页面……508
14.4.4 品牌页面……508
14.4.5 产品页面……514
14.4.6 搜索页面……515
14.5 亿赐客网站优化建议……516
14.5.1 涉及全站的调整……520
14.5.2 首页修改……533
14.5.3 一级分类页面……535
14.5.4 二级分类页面……538
14.5.5 三级分类页面（产品列表页面）……539
14.5.6 产品页面……542
14.5.7 产品按属性过滤页面……544
14.5.8 搜索页面……544
14.6 执行、效果及后续……545
附示 A SEO 术语……548

第 1 章

为什么要做 SEO

这将是非常简短的一章。如果您已经知道 SEO 对网站成功的意义，可以直接跳到第 2 章，开始学习 SEO 的具体方法。

1.1 什么是 SEO

SEO 是英文 Search Engine Optimization 的缩写，中文译为“搜索引擎优化”。简单地说，SEO 是指从自然搜索结果获得网站流量的技术和过程。复杂但更严谨些的定义如下：

SEO 是指在了解搜索引擎自然排名机制的基础上，对网站进行内部及外部的调整优化，改进网站在搜索引擎中的关键词自然排名，获得更多流量，从而达成网站销售及品牌建设的目标。

SEO 的完整意义和过程，随着读者阅读本书会越来越清晰。这里只对定义做简单说明。

在某种意义上看，SEO 是和搜索引擎博弈的过程。做 SEO，虽然不需要会编程，也不需要了解搜索引擎的技术细节，但理解搜索引擎的基本工作原理是必需的，不然只能是知其然，而不知其所以然，不能从根本上理解 SEO 技巧。了解搜索引擎原理，很多看似“新”的问题都可以迎刃而解。

网站的优化包括站内和站外两部分。站内优化指的是站长能控制的所有网站本身的调整，如网站结构、页面 HTML 代码。站外优化指的是外部链接建设及行业社群的参与互动，这些活动不是在网站本身进行的。

SEO 的研究对象是搜索引擎结果页面上的自然排名部分，通常在页面左侧，所以有百度左侧排名、Google 左侧排名等说法，与付费的搜索广告没有直接关系。

获得和提高关键词自然排名是 SEO 效果的表现之一，但最终目的是获得搜索流量，没有流量的排名是没有意义的。因此，关键词研究（针对有人搜索的关键词优化）、文案写作（吸引用户点击）十分重要。进一步说，SEO 追求的是目标流量，能最终带来赢利的流量。

网站的最终目标是完成转化，达到直接销售或品牌建设的目的。SEO、排名、流量都是手段。SEO 是网络营销的一部分，遇到与用户体验、业务流程等有冲突的情况时，一切以完成最多转化为最高原则，切不可为 SEO 而 SEO。

1.2 为什么要做 SEO

亲身做过网站的人都很清楚 SEO 的重要性。不排除有极小一部分网站不希望有人来看，比如我和太太给女儿写的博客，我们从不向别人提起，只有少数几个至亲好友知道，那几乎算是我们的私人日记，并不想让更多的人看到。但 99.9%的网站是希望有人来看的，而且人越多越好。不管网站赢利模式和目标是什么，有人来访问是前提。

而 SEO 是给网站带来访问者的最好方法，没有“之一”。

虽然也有其他网站推广的方法，运用得当时效果非凡。但总体来说，没有其他方法像 SEO 一样这么吸引人：

- 搜索流量质量高。其他方法是把网站推到用户眼前，用户本没有访问你的网站的意图。来自搜索的用户是在主动寻找你的网站，目标非常精准，转化率高。
- 性价比高。SEO 绝不是免费的，但确实是成本相对较低的，尤其是站长自己掌握 SEO 技术时。
- 可扩展性。只要掌握了关键词研究和内容扩展方法，网站可以不停地增加目标关键词及流量。
- 长期有效。网络广告、PPC（搜索广告）一旦停止投放，流量立即停止。事件营销效果明显，但话题过去，流量也就消失。而只要不作弊，搜索排名一旦上去，可以维持相当长时间，流量源源不断。
- 提高网站易用性，改善用户体验。SEO 是很少的（如果不是唯一的话）必须修改网站才能实现的推广方法之一，而 SEO 对页面的要求很多是与易用性相通的。

还没有亲手做过网站的读者，要明白为什么要做 SEO，其实也很简单，跟着下面这个逻辑：

- 近几年网络购物发展突飞猛进，非互联网圈子的人也知道网络是个刚刚开始被开采的金矿。
- 搜索引擎是网民寻找、比较、确定商品的最重要渠道，是电子商务的主要驱动力之一。
- 搜索引擎不是站长开的，不是你想排到前面就排到前面的。想办法把自己的网站排名提高，获得搜索流量，这就是 SEO。

根据艾瑞网的调查，2009 年中国网络购物交易规模为 2483.5 亿元，同比增长 93.7%。2007 年和 2008 年增长都在 100%以上，如图 1-1 所示。

网络购物占社会消费品零售总额的比例也在不断提高中，如图 1-2 所示。

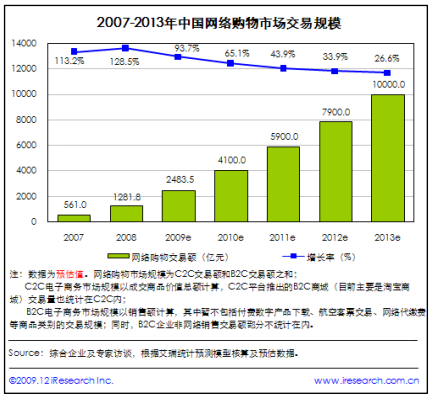


图 1-1　中国网购市场规模

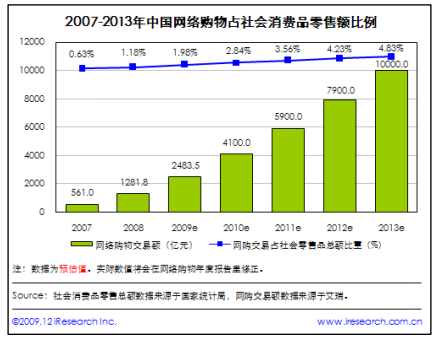


图 1-2　中国网购占社会消费品零售总额比例

这么大的网购规模代表了强劲需求，网民不在你的网站买，就得在别的网站买，谁进入用户视野，谁就赢得更多的付费客户。

那么，网民是怎样研究、比较产品和商家的呢？

中国互联网信息中心于 2010 年 7 月发布《第 26 次中国互联网络发展状况统计报告》，网民最常见的网络应用如图 1-3 所示。

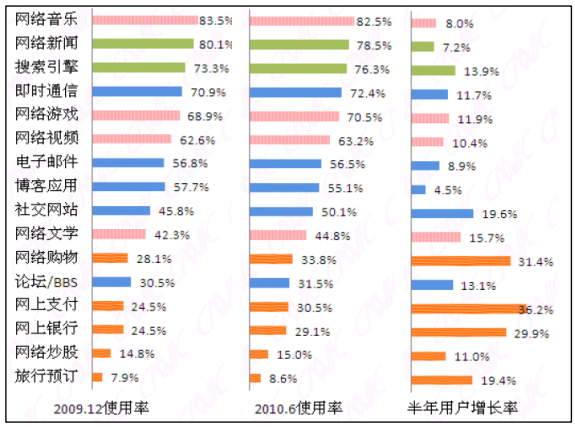


图 1-3　最常见的网络应用

除了听音乐、看新闻，用户最常用的就是搜索引擎，超过即时通信和电子邮件。从某种意义上说，搜索引擎已经成为一种习惯、一种生活方式，成为很多人获取信息的最重要方式。有了搜索引擎，我甚至都不用字典了，遇到记不清的单词，上 Google 输入印象中的拼法，Google 就会告诉我正确的拼法。

艾瑞网的统计也表明，中国搜索引擎用户覆盖率已经达到 80%以上，如图 1-4 所示。

根据互联网数据中心 DCCI 发布的报告，网民日常消费信息网上来源排在第二位的是搜索引擎，如图 1-5 所示。

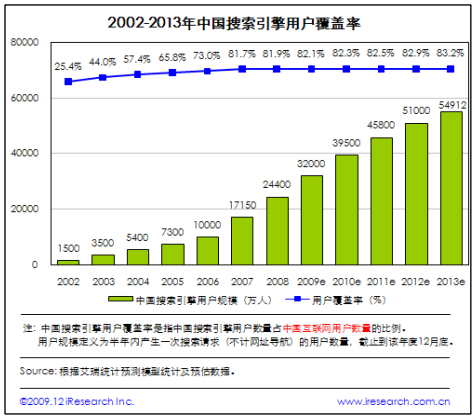


图 1-4　搜索引擎用户覆盖率

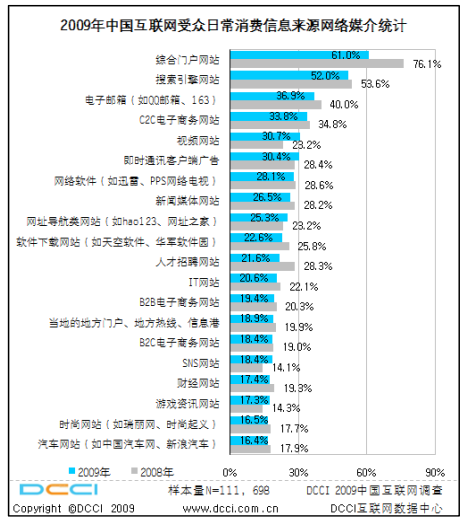


图 1-5　网民日常消费信息网上来源

国外的网络购物更是主要由搜索引擎驱动。网络分析公司 Compete 于 2010 年 2 月做了一项问卷调查，其中一个问题是“网上购物时你最常用哪个工具？”结果如图 1-6 和表 1-1 所示。

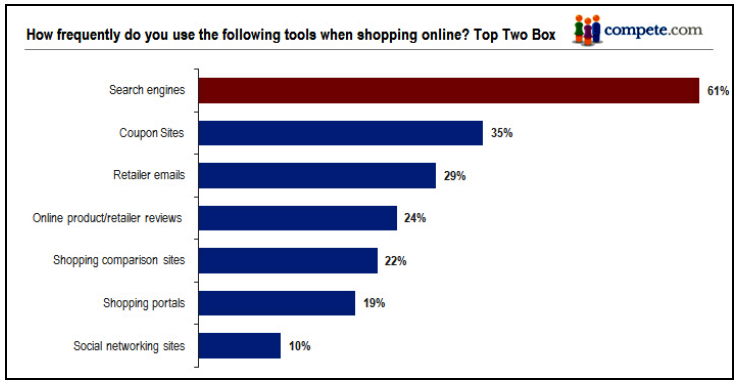


图 1-6　网上购物时最常用的工具

表 1-1　网上购物时最常用的工具

常 用 工 具	使　用　率
搜索引擎	61.00%
优惠券网站	35.00%
商家电子邮件	29.00%
线上产品/商家评测	24.00%
比较购物网站	22.00%
购物门户	19.00%
社会化网络网站	10.00%

搜索引擎是用户网购时最常用的工具，5 个人里就有 3 个说他们网上购物时总是或经常使用搜索引擎，远远超出其他工具或网站。对英文网站有了解的都知道，欧美用户更依赖搜索引擎，遇到任何问题，首先想到的就是用搜索引擎搜索。

越来越多的网站认识到搜索流量的重要性，这也体现在搜索引擎市场规模的快速增长上。2009 年中国搜索引擎市场规模达 69.5 亿元，相比 2008 年的 50.3 亿元，年同比增长 38.2%，如图 1-7 所示。

由于经济危机，2009 年增长已经算是缓慢的，2002 年到 2008 年的增长率都远远超过 38.2%。

搜索引擎占网络广告总规模比例也在不断增长中，这说明一部分广告主将预算更多地倾斜至性价比更高的搜索营销上，如图 1-8 所示。

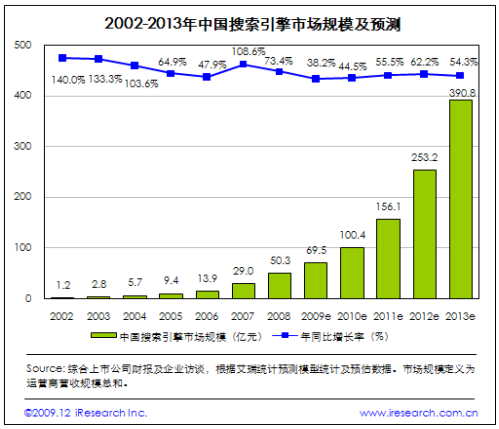


图 1-7　中国搜索引擎市场规模

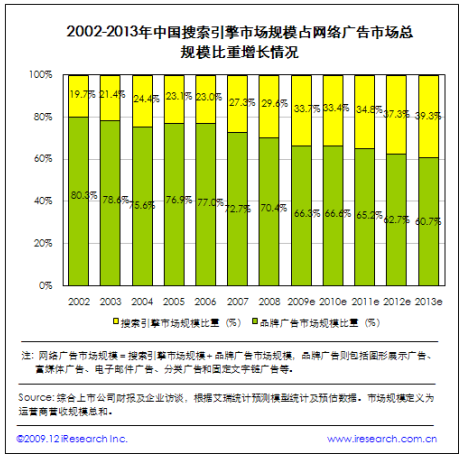


图 1-8　搜索引擎占网络广告市场比例

图 1-7、图 1-8 的搜索引擎市场规模指的是搜索广告（PPC）部分，而不是 SEO 的投入。搜索广告营收可以从搜索引擎和广告主公司获得较为准确的数字，但 SEO 的投入分散在大大小小的公司及个人站长上，又包含很多无形的投入，难认计算。

搜索引擎不仅驱动电子商务，对传统线下销售也有巨大影响。2007 年 7 月，雅虎和市场调查公司 comScore 发布了一项 2006 年 4 月至 2007 年 1 月做的跟踪调查，结果表明，搜索极大地促进了线下销售。当消费者接触到来自搜索的产品促销信息时，在线上每花 1 美元，就会在线下花 16 美元。而没有接触搜索信息的消费者，在线上每花 1 美元，会在线下花 6 美元。

相信很多读者也有这种经历：线上研究对比，线下购买。因此，就算是最传统的线下生意，无法进行线上销售，如果能被用户通过搜索引擎找到，也至关重要。

1.3 搜索引擎简史

随着互联网的发展，搜索引擎的出现是必然的。就像传统的图书馆一样，收藏的书籍、文件多了，必然出现管理、查找困难，索引和搜索就成为必需的了。实际上，搜索引擎原理在很大程度上源于传统文件检索技术。

网上资源数量远超出我们个人所能想象与掌控的，没有搜索引擎，我们几乎无法有效利用这些资源，也就没有互联网的今天。

当今社会化媒体如日中天，诸如 Facebook、Twitter 等应用的使用人次、网站流量、社会影响力已经达到甚至超过 Yahoo!、Google 等互联网巨人。有这样一种倾向和观点，网民现在寻找信息更多地是靠 Facebook、Twitter，而不是搜索引擎。也许搜索引擎已经过时了？

其实大可不必担心。也许 Google、百度会没落，会消失，但搜索引擎不会。就算依靠 Facebook、Twitter，或者以后出现的什么新网络服务寻找信息，用户要寻找信息时，一样要在搜索框输入关键词，这还是搜索，只不过搜索信息来源可能从搜索引擎收录的页面数据库变成 Facebook、Twitter 内部数据库，排名算法从页面相关性、链接变成会员、好友的推荐程度，但数据来源及算法的改变都不能改变用户对搜索的需要，也不会改变搜索的基本形式。

只要搜索存在，就有哪条信息排在前面的问题，就有 SEO 的存在。也许搜索引擎这个名字变了，SEO 也只是需要改个名字而已。

搜索是近 20 年互联网变化最快的领域之一，这种变化不仅体现在搜索技术突飞猛进和对网络经济的巨大推动上，搜索引擎本身的合纵连横、兴衰起伏也是精彩纷呈，常令人有眼花缭乱、瞠目结舌之感。

了解搜索引擎的发展历史有助于 SEO 人员理解搜索引擎营销的发展、变革，对未来有更准确的预期。本节就简要列出搜索引擎发展史上的重要事件，其中很多事件对今天搜索以及 SEO 行业的形态有至关重要的影响。

1990 年

第一个互联网上的搜索引擎 Archie 出现，用于搜索 FTP 服务器上的文件。当时基于 HTTP 协议的 Web 还没有出现。

1993 年

6 月，第一个 Web 搜索引擎 World Wide Web Wanderer 出现，它只收集网址，并没有索引文件内容。

10 月，第二个 Web 搜索引擎 ALIWEB 出现，开始索引文件元信息（也就是标题标签等信息），也还没有索引文件主体内容。

1994 年

1 月，Infoseek 创立，其搜索服务稍后才正式推出。Infoseek 是早期最重要的搜索引擎之一，允许站长提交网址是从 Infoseek 开始的。百度创始人李彦宏就是 Infoseek 的核心工程师之一。

4 月，Yahoo!由 David Filo 和 Jerry Yang（杨致远）创立。当时还没有注册 yahoo.com 域名，网站建在斯坦福大学的域名上。Yahoo!最初不是真正的搜索引擎，而是人工编辑的网站目录，创始人亲自把收集到的有价值的网站列在 Yahoo!目录中。当网站数量还不大时，可以实现人工编辑，既为用户提供了方便，又保证了信息质量，Yahoo!迅速成长为网络巨人。

4 月，第一个全文搜索引擎（索引文件全部内容）WebCrawler 推出。起初是华盛顿大学的一个研究项目。1995 年被美国在线 AOL 收购，1996 年又被 Excite 收购。2001 年停止研发自己的搜索技术，网站成为元搜索引擎（整合、显示多个第三方搜索引擎结果，被称为元搜索引擎）。

6 月，Lycos 创立，迅速成为最受欢迎的搜索引擎之一。

1995 年

1 月，yahoo.com 域名注册。4 月，Yahoo!公司正式成立。

12 月，Excite 搜索引擎正式上线，成为早期流行的搜索引擎之一。2001 年其母公司破产，被 InfoSpace 购买。2004 年被 Ask Jeeves 收购。

12 月，AltaVista 创立，迅速成为最受欢迎的搜索引擎，堪称当时的 Google。 AltaVista 在搜索领域做了很多开创性的工作，页面排名以站内因素为主。站长提交网址后收录非常迅速。

12 月，Infoseek 成为网景浏览器的默认搜索引擎。网景，Netscape，当时浏览器市场的绝对统治者，曾占市场份额的 90%以上。后来随着免费的微软 IE 浏览器推出而衰落，2008 年正式停止研发和技术支援。

1996 年

3 月，Google 作为创始人 Larry Page 和 Sergey Brin 在斯坦福大学的研究项目开始，当时的名称是 BackRub，1997 年才改名为 Google。

4 月，Yahoo!上市。

5 月，Inktomi 创立，是早期重要的搜索技术提供商，本身并没有供用户使用的搜索网站或界面，而是专门提供搜索技术给其他公司。Inktomi 最先开始付费收录，后来被 Google 等的成功证明行不通。

5 月，Hotbot 创立，开始时使用 Inktomi 数据。1998 年被 Lycos 收购，后来转型为元搜索引擎，显示来自 Google、FAST、Teoma 和 Inktomi 的结果。也是初期颇流行的搜索引擎之一。

11 月，Lycos 收录 6 千万文件，成为当时最大的搜索引擎。与今天的搜索引擎数据库相比，真是小巫见大巫。

1997 年

4 月，Ask Jeeves 上线，2006 年改名为 Ask，是唯一一个至今仍真实存在并有一定市场份额的早期搜索引擎。创立时标榜的特点是“自然语言”搜索，用户可以用问句形式搜索。起初雇用大量编辑进行人工编辑搜索结果，显然（至少今天看来很显然），这种模式行不通，后来不得不使用其他搜索引擎数据。

1998 年

2 月 21 日，GoTo（后来改名为 Overture）正式开始 Pay For Placement（出售搜索结果位置）服务，谁付的钱多，谁就排在前面。在当时饱受非议，是后来所有主流搜索引擎最主要收入来源 PPC 的始祖。

Direct Hit 创建，搜索结果的用户点击率是其排名因素的重要部分，Direct Hit 因此流行一时，但很快也因此造成搜索质量大幅下降。今天搜索引擎算法中引入用户浏览数据，早在初期搜索引擎就有探索并留下了教训，处理不好，就会被作弊者利用。

1998 年中，迪斯尼成为 Infoseek 的控股公司，将 Infoseek 转型为门户。早期搜索引擎没有明确赢利模式，遇到搜索质量下降、没有收入等困难时，常常采用转型为门户的方法，寄希望于像 Yahoo!那样赚网络显示广告的钱，但几乎没有成功转型的例子。这为 Google 等后来者坚持提高搜索技术、坚持简洁的搜索核心业务提供了前车之鉴。

同样在 1998 年，AltaVista 被卖给 Compaq，1999 年 10 月，Compaq 也将 AltaVista 转型为门户，AltaVista 走向没落。

1998 年中，Yahoo!放弃 1996 年开始使用的 AltaVista，转而使用 Inktomi 的搜索数据。早期 Yahoo!只在其目录中没有用户寻找的网站时才显示来自真正搜索引擎的数据。

9 月，Google 公司正式成立。

MSN 搜索推出，不过微软很长时间没有重视搜索引擎，直到 2004 年，MSN Search 都在使用其他提供商的搜索数据和技术。

1999 年

5 月，AllTheWeb.com 创建，是搜索技术公司 FAST 展示其技术的平台。

6 月，Netscape 放弃 Excite，开始使用 Google 搜索数据，对 Google 来说是个里程碑。

1999 年中，迪斯尼将 Infoseek 流量转入 Go.com，曾经流行一时的搜索引擎 Infoseek 消失。Go.com 几经波折，目前无声无息。

Lycos 停止自己的搜索技术，开始使用 AllTheWeb 数据。

2000 年

1 月，Ask Jeeves 以 5 亿美元收购搜索引擎 Direct Hit，但并没有进一步发展。2002 年年初 Direct Hit 正式结束。

1 月 18 日，百度成立。起初作为搜索技术提供商向其他网站提供中文搜索服务和数据。

5 月，Lycos 被西班牙公司 Terra Networks 收购，改名为 Terra Lycos。互联网泡沫破灭后，Terra Lycos 渐渐势微。

7 月，Yahoo!开始使用 Google 搜索数据，以自己当时最强大的网上品牌和流量培养出日后最强大的竞争对手。

2000 年中，GoTo.com 基本放弃用自己网站吸引用户的做法，开始向多家搜索引擎及网站提供付费搜索服务，包括 MetaCrawler.com，DogPile.com，Ask Jeeves，AOL，Netscape 等。这可能是 GoTo.com（后来的 Overture）的重大失误之一，一个具开创意义的模式，其品牌名称却没能被普通用户知道，也因此无法与 Google 这种家喻户晓的品牌抗衡。

10 月，Google 推出 AdWords，以 CPM 模式，也就是按显示付费，提供搜索广告服务。这个模式并不成功。

2001 年

9 月，Ask Jeeves 收购 Teoma，一个与 Google 一样重视链接的搜索引擎，曾经被认为是可能与 Google 抗衡的搜索引擎之一。

10 月，百度作为搜索引擎正式上线，直接独立提供搜索服务。中文搜索迅速进入百度时代。

2002 年

3 月，Google Adwords 推出 PPC 形式，也就是按点击是付费，成为今天 Adwords 的

主流。PPC 搜索广告由 Overture 发明，由 Google 发扬光大。从 PPC 开始，Google Adwords 才真正被客户接受和广泛使用，Google 成为充分利用搜索的网上赚钱机器。

5 月，美国在线 AOL 放弃 Inktomi ，开始使用 Google 搜索数据。

10 月，Yahoo!放弃在用户搜索时先返回 Yahoo!目录中数据的做法，完全改为显示 Google 搜索数据。Yahoo!目录至今还是最重要的网站目录。

12 月，Yahoo!收购苦于没有良好赢利模式的 Inktomi，为 2003 年一系列收购和整合拉开序幕。Yahoo!显然希望拥有自己的搜索技术，而不想继续依赖其他搜索引擎如 Google。

2003 年

2 月 18 日，Overture 宣布收购 AltaVista 。除了广告平台，Overture 也拥有了自己的搜索技术。

2 月 25 日，Overture 宣布收购 FAST 的搜索技术部门，FAST 拥有 AllTheWeb.com，也向另一个著名的搜索网站 Lycos 提供搜索数据。Overture 因此拥有了当时两大主要搜索技术公司。

3 月，Google 推出后来被称为 Adsense 的内容广告系统，向其他内容网站提供广告服务，成为很多网站的主要收入来源之一。

7 月，Yahoo!宣布以 16 亿美元收购 Overture，将除了 Google 之外的几乎所有主流搜索技术（Inktomi 、AltaVista、FAST）收归旗下。可惜，随着 2010 年 Yahoo！放弃自己的搜索技术，转而使用微软必应，Yahoo!以前收购的及自己在此基础上研发多年的搜索技术全部无疾而终。Overture 的 PPC 广告平台被整合、改名为 Yahoo! Search Marketing。

2003 年，微软 MSN 开始开发自己的搜索引擎技术。此前，MSN 网站一直使用 Inktomi 等搜索技术提供商的搜索数据。

2004 年

2 月，正如所有人预料的，Yahoo!在收购了几大搜索公司后推出自己的搜索引擎，不再使用 Google 数据和技术。

8 月，Google 上市。

11 月，微软推出自己的搜索引擎 MSN Search，不再使用第三方搜索服务。三雄鼎立时期开始。

2005 年

8 月，百度上市。

2006 年

5 月，微软推出类似于 Google Adwords 的广告系统 AdCenter 。

9 月，MSN Search 改名 Live Search，实际上，MSN 网络品牌全部改为 Live，并做了大量推广。不过并没有显著提高微软在搜索市场的份额，Google 仍然一枝独秀。

2007 年

3 月，Google 也开始提供类似网站联盟的按转化付费的广告形式。4 月，Google 收购传统网络广告公司 DoubleClick，进入更广泛的网络广告领域。

2009 年

6 月，微软 Live Search 改名为必应（Bing）。

7 月 29 日，微软和 Yahoo!达成历史性协议，Yahoo!将逐步放弃自己的搜索技术，使用 Bing 数据。此前几年，Yahoo!始终挣扎在困境中，CEO 几度更换，大股东内讧，搜索市场份额不断下降。终于，Yahoo!被自己亲手培养的 Google 彻底打败在搜索战场。Yahoo!曾经是英雄，但 Yahoo!时代正式结束了。

2010 年

8 月 25 日，Yahoo!开始使用 Bing 搜索数据。

2011 年

2 月 24 日，Google 推出旨在减少搜索结果中低质量页面的 Panda 更新，直至 11 月还在不定期推出微调。Panda 更新对 SEO 业界影响深远。

2003 年至 2011 年，搜索领域的技术革新不断，以 Google 为代表的搜索引擎推出了整合搜索、个人化搜索、实时搜索、地图服务、线上文件编辑、网站统计、浏览器、网管工具、超大容量电子邮件、即时通信等。总体上看，通常是 Google 推出新服务，其他搜索引擎很快跟进。

国内，百度也创建了百科、知道、贴吧等独特服务。

但 2009 年以前搜索引擎服务商没有大的变化，国际上始终是 Google 独大，Yahoo!第二，占有不大不小的份额，微软 Live/Bing 第三，苦苦追赶而不得。2009 年微软推出 Bing，Yahoo!开始使用 Bing 技术后，严格地说，Yahoo!已自动退出搜索引擎市场。

据 NetMarketShare 统计，2010 年 8 月全球搜索引擎市场份额分布如图 1-9 所示。

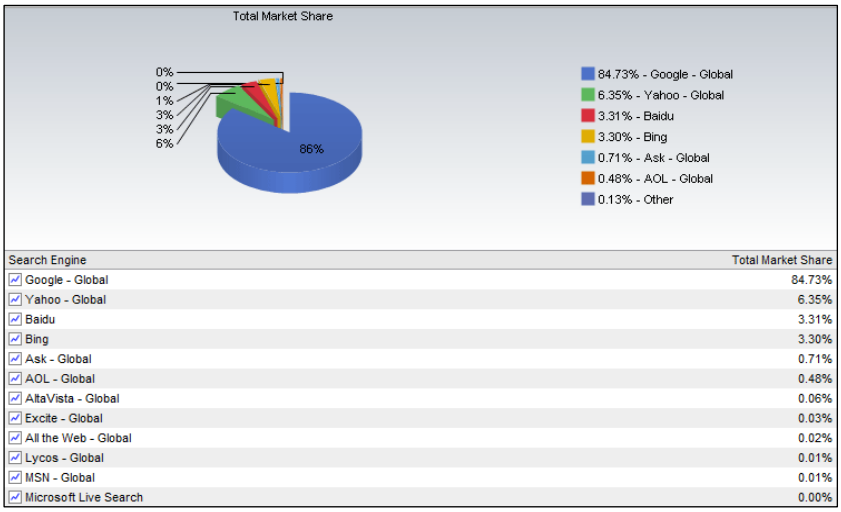


图 1-9　2010 年 8 月全球搜索引擎市场份额

Google 占据 84.73%，居于绝对领先地位。由于中国用户数的庞大，百度搜索量排在第三位，但除中国外，其他国家很少使用百度。

美国市场一直是全球搜索引擎风向标。据 HitWise 统计数字，2010 年 1 至 8 月各搜索引擎市场份额变化情况如表 1-2 所示。

表 1-2　2010 年美国搜索引擎市场份额

数据分布日期	Google	Yahoo!	Bing	Ask	AOL
8 月 28 日	71.59%	14.28%	9.87%	2.28%	1.21%
7 月 31 日	71.43%	14.43%	9.86%	2.32%	1.19%
6 月 26 日	71.65%	14.37%	9.85%	2.19%	1.15%
5 月 22 日	72.00%	14.58%	9.20%	2.18%	1.06%
5 月 8 日	71.56%	14.79%	9.31%	2.27%	1.07%
3 月 6 日	71.07%	14.46%	9.55%	3.01%	0.98%
2 月 6 日	71.35%	14.60%	9.56%	2.55%	1.06%
1 月 2 日	72.25%	14.83%	8.91%	2.53%	0.77%

不同的市场调查公司采样不同，计算方法不同，得出的数据也有差别。同样是美国 2010 年 8 月数字，Nielsen 公司的数据是 Google 占 65%，Bing 已经超过 Yahoo，以 13.9% 的市场份额上升为第二名，Yahoo 为 13.1% 。而 comScore 的报告又显示 2010 年 8 月 Google 在美国的市场份额为 65.4%，呈下降趋势，Yahoo 为 17.4%，比前几个月还有增长，Bing 为 11.1%，也略有增长。

2011 年 10 月的 HitWise 数据显示，Google 占有美国搜索市场的 65.38%，比一年前有了小幅但明显的降低。Bing（包括 Yahoo!搜索）市场份额为 28.62%。2011 年 11 月

comScore 数据显示，Google 在美国搜索市场占 65.4%份额，Bing 占 15%，雅虎占 15.10%。

在中文搜索市场，百度独大，Google 第二。2010 年，Google 退出中国风波后，Google 在中文搜索市场份额持续下降。

据艾瑞网公布的数字，2010 年第二季度，按网页搜索量计算，百度在中国搜索引擎市场占有 80.2%份额，Google 下降至 14.1%，如图 1-10 所示。

按营收计算，百度占 70.8%，Google 占 27.3%。Google 每次搜索产生的收入高于其他搜索引擎，如图 1-11 所示。

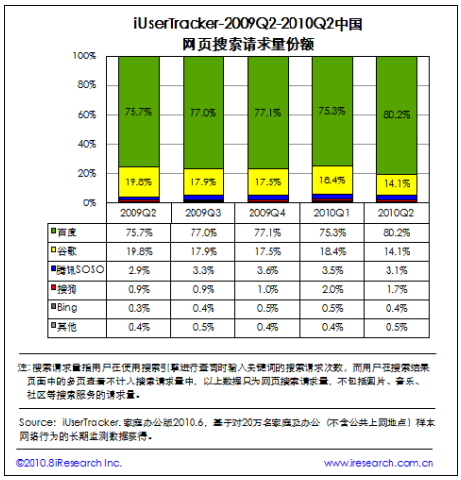


图 1-10　2010 年中国搜索市场份额

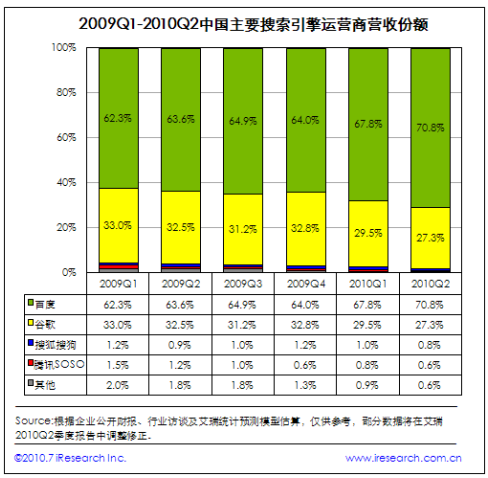


图 1-11　2010 年中国搜索引擎市场营收份额

搜索领域变化很快。就在本书第一版最终完稿时（2010 年 9 月），百度推出了开放平台，用户不离开百度搜索结果页面就可以使用某些应用（如小游戏）。Google 推出了“Google Instant”，用户正在输入搜索词时（还没有按搜索按钮或 Enter 键），就可以看到搜索结果。这些都将影响用户搜索模式，对 SEO 有什么长远影响还有待观察。

第 2 章
了解搜索引擎

一个合格的 SEO 必须了解搜索引擎基本工作原理。很多看似令人迷惑的 SEO 原理及技巧，其实从搜索引擎原理出发，都是自然而然的事情。

为什么要了解搜索引擎原理？

说到底，SEO 是在保证用户体验的基础上尽量迎合搜索引擎。与研究用户界面及可用性不同的是，SEO 既要从用户出发，也要站在搜索引擎的角度考虑问题，才能清晰地知道怎样优化网站。SEO 人员必须知道：搜索引擎要解决什么问题，有哪些技术上的困难，有什么限制，搜索引擎又怎样取舍。

从某个角度来说，SEO 人员优化网站就是尽量减少搜索引擎的工作量、降低搜索引擎的工作难度，使搜索引擎能更轻松、快速地收录网站页面，更准确地提取页面内容。不了解搜索引擎工作原理，也就无从替搜索引擎解决一些 SEOer 力所能及的技术问题。当搜索引擎面对一个网站，发现要处理的问题太多、难度太大时，搜索引擎可能就对这样的网站敬而远之 了。

很多 SEO 技巧是基于对搜索引擎的理解。下面举几个例子。

我们都知道网站域名和页面权重非常重要，这是知其然，很多人不一定知其所以然。权重除了意味着权威度高、内容可靠，因而容易获得好排名外，获得一个最基本的权重，也是页面能参与相关性计算的最基本条件。一些权重太低的页面，就算有很高的相关性也很可能无法获得排名，因为根本没有机会参与排名。

比如很多 SEOer 津津乐道的“伪原创”。首先，抄袭是不道德甚至违法的行为，把别人的文章拿来加一些“的、地、得”，段落换换顺序就当成自己的原创放在网站上，这是令人鄙视的抄袭行为。理解搜索引擎原理的话，就会知道这样的伪原创也不管用。搜索引擎并不会因为两篇文章差几个字、段落顺序不同，就真的把它们当成不同的内容。搜索引擎的权重算法要先进、准确得多。

再如，对大型网站来说，最关键的问题是解决收录。只有收录充分，才能带动大量长尾关键词。就算是有人力、财力的大公司，当面对几百万几千万页面的网站时，也不容易处理好充分收录的问题。只有在深入了解搜索引擎蜘蛛爬行原理的基础上，才能尽量使蜘蛛抓得快而全面。

上面所举的几个例子，读者看完搜索引擎原理简介这一节后，会有更深入的认识。

2.1 搜索引擎与目录

早期的 SEO 资料经常把真正的搜索引擎与目录放在一起讨论，甚至把目录也称为搜索引擎的一种，这种讲法并不准确。

真正的搜索引擎指的是由蜘蛛程序沿着链接爬行和抓取网上的大量页面，存进数据库，经过预处理，用户在搜索框输入关键词后，搜索引擎排序程序从数据库中挑选出符合搜索关键词要求的页面。蜘蛛的爬行、页面的收录及排序都是自动处理。

网站目录则是一套人工编辑的分类目录，由编辑人员人工创建多个层次的分类，站长可以在不同分类里提交网站，目录编辑在后台审核所提交的网站，将网站放置于相应的分类页面。有的时候编辑也主动收录网站。典型的网站目录包括雅虎目录、开放目录、hao123 等。

目录并不是本书中所讨论的 SEO 所关注的真正的搜索引擎。虽然网站目录也常有一个搜索框，但目录的数据来源是人工编辑得到的。

搜索引擎和目录两者各有优劣。

搜索引擎收录的页面数远远高于目录能收录的页面数。但搜索引擎收录的页面质量参差不齐，对网站内容和关键词提取的准确性通常也没有目录高。

限于人力，目录能收录的通常只是网站首页，而且规模十分有限，不过收录的网站通常质量比较高。像雅虎、开放目录、hao123 这些大型目录，收录标准非常高。目录收录网站时存储的页面标题、说明文字都是人工编辑的，比较准确。

搜索引擎数据更新快，而目录中收录的很多网站内容十分陈旧，甚至网站可能已经不再存在了。

雅虎目录、搜狐目录等曾经是用户在网上寻找信息的主流方式，给用户的感觉与真正的搜索引擎也相差不多。这也就是目录有时候被误称为“搜索引擎的一种”的原因。但随着 Google 等真正意义上的搜索引擎发展起来，目录的使用迅速减少，现在已经很少有人使用网站目录寻找信息了。现在的网站目录对 SEO 的最大意义是建设外部链接，像雅虎、开放目录、hao123 等都有很高的权重，可以给被收录的网站带来一个高质量的外部链接。

2.2 搜索引擎面对的挑战

搜索引擎系统是最复杂的计算系统之一，当今主流搜索引擎服务商都是有财力、人力的大公司。即使有技术、人力、财力的保证，搜索引擎还是面临很多技术挑战。搜索引擎诞生后的十多年中，技术已经得到了长足的进步。我们今天看到的搜索结果质量与

10 年前相比已经好得多了。不过这还只是一个开始，搜索引擎必然还会有更多创新，提供更多、更准确的内容。

总体来说，搜索引擎主要会面对以下几方面的挑战。

1．页面抓取需要快而全面

互联网是一个动态的内容网络，每天有无数页面被更新、创建，无数用户在网站上发布内容、沟通联系。要返回最有用的内容，搜索引擎就要抓取最新的页面。但是由于页面数量巨大，搜索引擎蜘蛛更新一次数据库中的页面要花很长时间。搜索引擎刚诞生时，这个抓取周期往往以几个月计算。这也就是 Google 在 2003 年以前每个月有一次大更新的原因所在。

现在主流搜索引擎都已经能在几天之内更新重要页面，权重高的网站上的新文件几小时甚至几分钟之内就会被收录。不过，这种快速收录和更新也只能局限于高权重网站。很多页面几个月不被重新抓取和更新，也是常见的。

要返回最好的结果，搜索引擎也必须抓取尽量全面的页面，这就需要解决很多技术问题。一些网站并不利于搜索引擎蜘蛛爬行和抓取，诸如网站链接结构的缺陷、大量使用 Flash、JavaScript 脚本，或者把内容放在用户必须登录以后才能访问的部分，都增大了搜索引擎抓取内容的难度。

2．海量数据存储

一些大型网站单是一个网站就有百万千万个页面，可以想象网上所有网站的页面加起来是一个什么数据量。搜索引擎蜘蛛抓取页面后，还必须有效存储这些数据，数据结构必须合理，具备极高的扩展性，写入及访问速度要求也很高。

除了页面数据，搜索引擎还需要存储页面之间的链接关系及大量历史数据，这样的数据量是用户无法想象的。据说 Google 有几十个数据中心，上百万台服务器。这样大规模的数据存储和访问必然存在很多技术挑战。

我们经常在搜索结果中看到，排名会没有明显原因地上下波动，甚至可能刷新一下页面，就看到不同的排名，有的时候网站数据也可能丢失。这些都可能与大规模数据存储的技术难题有关。

3．索引处理快速有效，具可扩展性

搜索引擎将页面数据抓取和存储后，还要进行索引处理，包括链接关系的计算、正向索引、倒排索引等。由于数据库中页面数量大，进行 PR 之类的迭代计算也是耗时费力的。要想及时提供相关又及时的搜索结果，仅仅抓取没有用，还必须进行大量索引计算。由于随时都有新数据、新页面加入，因此索引处理也要具备很好的扩展性。

4．查询处理快速准确

查询是普通用户唯一能看到的搜索引擎工作步骤。用户在搜索框输入关键词，单击“搜索”按钮后，通常不到一秒就会看到搜索结果。表面最简单的过程，实际上涉及非常复杂的后台处理。在最后的查询阶段，最重要的难题是怎样在不到一秒的时间内，快速从几十万、几百万，甚至几千万包含搜索词的页面中，找到最合理、最相关的 1000 个页面，并且按照相关性、权威性排列。

5．判断用户意图及人工智能

应该说前 4 个挑战现在的搜索引擎都已经能够比较好地解决，但判断用户意图还处在初级阶段。不同用户搜索相同的关键词，很可能是在寻找不同的东西。比如搜索“苹果”，用户到底是想了解苹果这个水果，还是苹果电脑？还是电影《苹果》的信息？没有上下文，没有对用户个人搜索习惯的了解，就完全无从判断。

搜索引擎目前正在致力于基于用户搜索习惯及历史数据的了解上，判断搜索意图，返回更相关的结果。今后搜索引擎是否能达到人工智能水平，真正了解用户搜索词的意义和目的，让我们拭目以待。

2.3 搜索结果显示格式

让我们先稍微深入地了解一下搜索结果的展现形式。

2.3.1 搜索结果页面

用户在搜索引擎搜索框中输入关键词，单击“搜索”按钮后，搜索引擎在很短时间内返回一个搜索结果页面。如图 2-1 所示是 Google 的搜索结果页面，也是比较典型的搜索结果页面排版格式。

页面主体有两部分最主要，一是广告，二是自然搜索结果。如图 2-1 所示，页面右侧 8 个结果及左侧最上面的一个结果，都标注为“赞助商链接”，这就是广告。绝大部分网民都比较清楚右侧显示的是广告，所以右侧赞助商链接没有加特殊底色。页面左侧上部的广告链接使用浅黄色底色，可以和下面的自然搜索结果清楚地分开。右侧广告最多有 8 个，上部广告可以多至 3 个。

搜索广告在网络营销行业经常称为 PPC，由广告商针对关键词进行竞价，广告显示广告商无须付费，只有搜索用户点击广告后，广告商才按竞价价格支付广告费用。PPC 是搜索营销的另一个主要内容。

搜索结果页面左侧广告下面，占据页面最大部分的就是自然搜索结果。通常每个页面会列出 10 个自然搜索结果。用户可以在账户设置中选择每页显示 100 个搜索结果。每

个搜索结果的格式后面再做介绍。

图 2-1　Google 的搜索结果页面

页面最左上角是垂直搜索链接，用户点击后可以直接访问图片、视频、地图等搜索结果。

搜索框右下方显示满足搜索关键词的结果总数，例如图 2-1 中所显示的 19 600 000 条结果。这个搜索结果数是研究竞争程度的依据之一。

自然搜索结果下面显示相关搜索。搜索引擎根据用户搜索数据，列出相关的其他搜索词。

页面最下面又是一个赞助商广告，与页面顶部的广告相同。页面左侧顶部及左侧底部的广告，并不是每次搜索有广告商竞价时都会出现，只有点击率和质量分数达到一定水平的广告才会出现在左侧顶部或底部。

SEOer 最关注的是占据页面主体的自然搜索结果。统计数据显示，自然搜索结果总

点击访问数要远远大于广告点击数。但是企业花费在 SEO 上的费用却远远低于花费在搜索广告上的费用。这既是 SEO 的尴尬，也是最大的机会。掌握了 SEO 流量，才能掌握最大搜索流量。

我们再来看百度搜索结果页面，如图 2-2 所示。

图 2-2　百度的搜索结果页面

百度搜索结果页面与 Google 大致相同，区别在于广告部分的显示方法。如图 2-1 所示，右侧也是最多 8 个广告，不过并没有标注为赞助商链接或加其他提示文字。左侧最上面标注为“推广链接”的结果也是广告，这是百度启用凤巢系统后显示的广告。不过这几个广告只加了非常浅的灰色背景，不注意看几乎无法与下面的自然搜索结果区分开来。

有的关键词搜索没有触发凤巢系统广告，还会继续显示传统百度广告，如图 2-3 所示。

图 2-3　传统百度搜索广告

传统百度左侧广告既没有明确标注为推广链接，也没有使用任何背景颜色，与下面的自然搜索结果更不容易分辨。唯一能分辨出上面 3 个是广告的，是结果列表最后一行最右侧标有“推广”两个字。百度广告结果在背景颜色、文字标注上，都比较难以与自

然结果区分。SEO 人员当然很清楚这两者之间的区别，普通网民却难以察觉，尤其是百度传统左侧广告。这也是百度搜索结果常为人诟病的原因之一。

2.3.2 经典搜索结果列表

我们再来看看每一个搜索结果页面的展现格式。如图 2-4 所示是百度的搜索结果列表，主要分三部分。

SEO每天一贴 – Zac的SEO博客，正在写书，暂时每月两三贴
SEO每天一贴研究搜索引擎优化SEO技术,网络营销及电子商务思考。正在写一本SEO方面的书,所以不能每天一贴了,会尽快恢复。
www.chinamyhosting.com/seoblog/ 2010-2-24 - 百度快照

图 2-4 最常见的搜索结果列表格式

第一行是页面标题，通常取自页面 HTML 代码中的标题标签（Title Tag）。这是结果列表中最醒目的部分，用户点击标题就可以访问对应的网页。可见页面标题标签的写法，无论对排名还是对点击率都有重要意义。

第二行、第三行是页面说明。页面说明有的时候取自页面 HTML 中的说明标签（Description Tag），有时从页面可见文字中动态抓取相关内容。显示什么页面说明文字是用户查询时才决定的。

某些与日期有明确联系的页面，Google 会在说明文字最前面显示日期，省略号后再显示页面说明。如博客帖子这类有明确发布日期的页面，如图 2-5 所示。

URL网址规范化– 搜索引擎优化SEO每天一贴
2006年4月10日 ... URL网址规范化- 作者:Zac，首发于搜索引擎优化SEO每天一贴。
www.chinamyhosting.com/.../04/.../url-canonicalization/ - 网页快照 - 类似结果

图 2-5 Google 搜索结果列表中显示日期

第四行显示三个信息。最左侧是网址，用户可以看到页面来自哪个网站，以及目录、文件名信息。

中间是百度数据库中页面最后更新的日期。

然后是百度快照链接，用户可以点击快照，查看存储在百度数据库中的页面内容。当页面被删除或者有其他技术问题导致不能打开网站时，用户至少还可以从快照中查看想要的内容。

用户所搜索的关键词在标题及说明部分都用红色高亮显示。用户可以非常快速地看到页面与自己搜索的关键词相关性如何。如图 2-6 中的“SEO”这三个字母。

SEO每天一贴– Zac的SEO博客，正在写书，暂时每月两三贴
SEO每天一贴研究搜索引擎优化SEO技术，网络营销及电子商务思考。正在写一本SEO方面的书，所以不能每天一贴了，会尽快恢复。
www.chinamyhosting.com/seoblog/ - 9 小时前 - 网页快照 - 类似结果

图 2-6 搜索词在 Google 中文搜索结果列表中高亮显示

Google 结果列表与百度大致相同，几处小的区别如下：

搜索关键词在 URL 中出现时加粗显示，如图 2-6 中的 SEO 三个字母。

URL 右侧的页面最近更新时间不是按日期显示的，而是显示为几小时前。

网页快照链接后面有时还有一个“类似结果”链接。用户点击类似结果后可以看到与这个页面相似的其他网页。

Google 英文结果列表与中文还有一个区别，就是搜索关键词在标题及说明部分不是红色高亮显示的，而是加粗显示的，如图 2-7 所示。

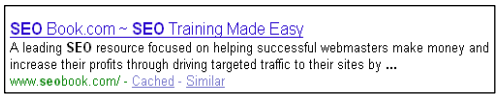


图 2-7　搜索词在 Google 英文搜索结果列表中加粗显示

红色高亮显示应该是 Google 为适应中国用户搜索习惯所做的变化。

2010 年 4 月，本章初稿完成后，笔者又发现百度也在试验将 URL 中的关键词加粗，但不是所有 URL 中的关键词都会加粗，如图 2-8 所示。

图 2-8　搜索词在百度列表中 URL 目录部分加粗显示

上面讨论的是最经典的结果列表格式。搜索引擎近几年也在不停尝试不同格式的搜索结果格式，尤其是 Google，推出了很多特色结果列表，下面选几个主要的进行简单介绍。

2.3.3　整合搜索结果

在前面的 Google 搜索结果页面抓图中，大家就能看到两个整合搜索结果：中间的资讯结果和页面底部的博客搜索结果。根据搜索关键词的不同，Google 还经常把其他垂直搜索结果混合在正常网页搜索结果中，比如图片结果、视频结果，如图 2-9 和图 2-10 所示。

图 2-9　图片搜索结果

图 2-10　视频搜索结果

百度也有类似的整合搜索结果，主要以新闻内容为主，其他则很少见到，如图 2-11 所示。

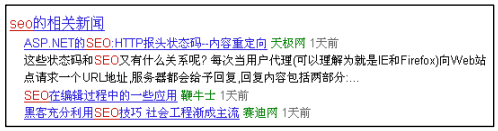


图 2-11　百度的新闻搜索结果

2.3.4　缩进列表

缩进列表（Indented Listing）是指当搜索结果页面上应该出现两个来自同一网站的页面时，比如第一位和第八位，按照正常排名算法是来自同一网站的两个页面，Google 不是把它们排在第一和第八位，而是把两个结果连在一起排在第一和第二位，第二位的结果向右侧缩进三个字的空间，如图 2-12 所示。这样的列表格式非常有助于提高点击率。

图 2-12　缩进列表

2.3.5　全站链接

对某些权重比较高的网站，当用户搜索一个关键词，这个网站的结果是最权威的内容来源时，Google 除了正常结果列表外，还会显示最多四行、两列共 8 个内页链接，称为全站链接（Sitelinks），如图 2-13 所示。

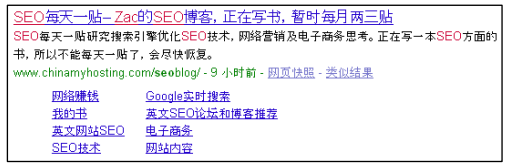


图 2-13　全站链接

这无疑为权重高的网站提供了多几倍的访问入口，视觉上的醒目也大大提高了点击率。

2.3.6 迷你全站链接

权重高的网站在某些情况下还会显示迷你全站链接（Mini Sitelinks），不是 4 行 8 个，而是 1 行 4 个链接。显示的内页与上面说的全站链接是一样的，取其中前面 4 个，如图 2-14 所示。

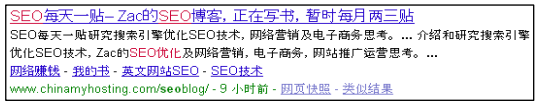


图 2-14　迷你全站链接

2.3.7 One-box

某些关键词会触发 Google One-box 结果，直接在搜索结果页面上显示相关信息，用户不用点击到其他网站上查看。如图 2-15 显示搜索“北京银行”时显示的股价 One-box。

图 2-15　Google 的 One-box

2.3.8 富摘要

某些使用 RDFa 或 Microdata 格式标签的页面，Google 可能还会在标题下面以灰色文字加一行富摘要（Rich Snippet），如图 2-16 所示的论坛帖子页面还显示出帖子个数、作者数及更新日期。

这样的排版格式无疑也会提高关注度和点击率。在富摘要中显示合适的信息，有助于说服用户点击结果，比如显示产品价格、用户评分、用户评论数目等。

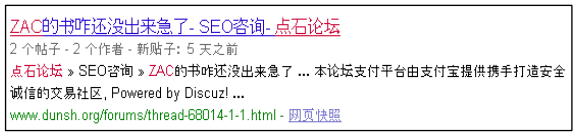


图 2-16　Google 的富摘要

百度也有类似显示方式，如图 2-17 所示。

2007年点石最后一次线下活动南京茶话会专题 - 点石线下茶话会 - ...
14个回复 - 发帖时间: 2007年11月28日
点石论坛...点石茶话会一直秉承交流分享的宗旨，会在中国SEO技术爱好者和从业人员比较密集... 17：00~17：50 互动问答 本环节由演讲嘉宾帮助大家回答疑问，请参会...
www.dunsh.org/forums/thread-16295-1-1.html 2010-3-29 - 百度快照

图 2-17　百度的富摘要

2.3.9　面包屑导航

Google 最近又在结果列表中大规模使用面包屑导航。原本显示一个网址的地方，改为面包屑导航格式，其中的每一个分类链接都指向网站上相应的分类页面，如图 2-18 所示。

药物减肥_39健康减肥_39健康网
介绍各种减肥产品，减肥药品知识，服用减肥药的注意事项，减肥药减肥的基本原理，减肥药的主要成分及副作用等。
fitness.39.net › 减肥 › 减肥方法 - 网页快照 - 类似结果

图 2-18　Google 显示的面包屑导航

用户不仅可以点击标题访问产品页面，还可以直接从搜索列表的面包屑导航中点击上级分类链接访问分类页面。

2.3.10　说明文字中的链接

对一部分使用了页面内锚链接的页面，Google 有时也尝试在说明文字中显示链接，用户可以跳到页内锚链接部分，如图 2-19 所示。

Jeff Dean-程序员百科全书- 科技中国——欢迎光临全球最大的互联网博物馆 ...
跳到谷歌自**建数据中心**基础设施: 2009-06-29: 了Google一向很少对外透露其数据中心的工作，但5 .... 一窥Google数据中心自行定制的40台服务器机柜。...
www.techcn.com.cn › 科技人物百科全书 - 网页快照 - 类似结果

图 2-19　Google 显示的说明文字中的链接

这种显示方式目前还比较少见。

上面介绍的一些搜索列表变化形式有逐渐增多的趋势，不过它们的基本形式与经典搜索结果列表相差不大，最经典的结果列表还是最常见的。

2.4　搜索引擎工作原理简介

搜索引擎工作过程非常复杂，接下来的几节我们简单介绍搜索引擎是怎样实现网页排名的。这里介绍的内容相对于真正的搜索引擎技术来说只是皮毛，不过对 SEO 人员已经足够用了。

搜索引擎的工作过程大体上可以分成三个阶段。

（1）爬行和抓取：搜索引擎蜘蛛通过跟踪链接访问网页，获得页面 HTML 代码存入数据库。

（2）预处理：索引程序对抓取来的页面数据进行文字提取、中文分词、索引等处理，以备排名程序调用。

（3）排名：用户输入关键词后，排名程序调用索引库数据，计算相关性，然后按一定格式生成搜索结果页面。

2.4.1 爬行和抓取

爬行和抓取是搜索引擎工作的第一步，完成数据收集的任务。

1．蜘蛛

搜索引擎用来爬行和访问页面的程序被称为蜘蛛（spider），也称为机器人（bot）。

搜索引擎蜘蛛访问网站页面时类似于普通用户使用的浏览器。蜘蛛程序发出页面访问请求后，服务器返回 HTML 代码，蜘蛛程序把收到的代码存入原始页面数据库。搜索引擎为了提高爬行和抓取速度，都使用多个蜘蛛并发分布爬行。

蜘蛛访问任何一个网站时，都会先访问网站根目录下的robots.txt文件。如果robots.txt文件禁止搜索引擎抓取某些文件或目录，蜘蛛将遵守协议，不抓取被禁止的网址。

和浏览器一样，搜索引擎蜘蛛也有标明自己身份的代理名称，站长可以在日志文件中看到搜索引擎的特定代理名称，从而辨识搜索引擎蜘蛛。下面列出常见的搜索引擎蜘蛛名称：

- Baiduspider+(+http://www.baidu.com/search/spider.htm) 百度蜘蛛
- Mozilla/5.0 (compatible; Yahoo! Slurp China; http://misc.yahoo.com.cn/ help.html) 雅虎中国蜘蛛
- Mozilla/5.0 (compatible; Yahoo! Slurp/3.0; http://help.yahoo.com/help/us/ysearch/slurp) 英文雅虎蜘蛛
- Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html) Google 蜘蛛
- msnbot/1.1 (+http://search.msn.com/msnbot.htm) 微软 Bing 蜘蛛
- Sogou+web+robot+(+http://www.sogou.com/docs/help/webmasters.htm#07)搜狗蜘蛛
- Sosospider+(+http://help.soso.com/webspider.htm) 搜搜蜘蛛
- Mozilla/5.0 (compatible; YodaoBot/1.0; http://www.yodao.com/help/webmaster/spider/;) 有道蜘蛛

2．跟踪链接

为了抓取网上尽量多的页面，搜索引擎蜘蛛会跟踪页面上的链接，从一个页面爬到下一个页面，就好像蜘蛛在蜘蛛网上爬行那样，这也就是搜索引擎蜘蛛这个名称的由来。

整个互联网是由相互链接的网站及页面组成的。从理论上说，蜘蛛从任何一个页面出发，顺着链接都可以爬行到网上的所有页面。当然，由于网站及页面链接结构异常复杂，蜘蛛需要采取一定的爬行策略才能遍历网上所有页面。

最简单的爬行遍历策略分为两种，一种是深度优先，另一种是广度优先。

所谓深度优先，指的是蜘蛛沿着发现的链接一直向前爬行，直到前面再也没有其他链接，然后返回到第一个页面，沿着另一个链接再一直往前爬行。

如图 2-20 所示，蜘蛛跟踪链接，从 A 页面爬行到 A1，A2，A3，A4，到 A4 页面后，已经没有其他链接可以跟踪就返回 A 页面，顺着页面上的另一个链接，爬行到 B1，B2，B3，B4。在深度优先策略中，蜘蛛一直爬到无法再向前，才返回爬另一条线。

广度优先是指蜘蛛在一个页面上发现多个链接时，不是顺着一个链接一直向前，而是把页面上所有第一层链接都爬一遍，然后再沿着第二层页面上发现的链接爬向第三层页面。

如图 2-21 所示，蜘蛛从 A 页面顺着链接爬行到 A1，B1，C1 页面，直到 A 页面上的所有链接都爬行完，然后再从 A1 页面发现的下一层链接，爬行到 A2，A3，A4，……页面。

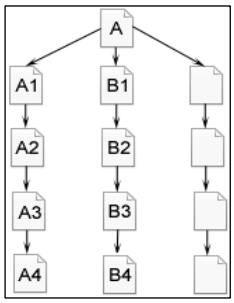


图 2-20　深度优先遍历策略

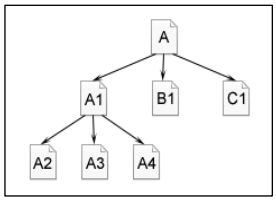


图 2-21　广度优先遍历策略

从理论上说，无论是深度优先还是广度优先，只要给蜘蛛足够的时间，都能爬完整个互联网。在实际工作中，蜘蛛的带宽资源、时间都不是无限的，也不可能爬完所有页面。实际上最大的搜索引擎也只是爬行和收录了互联网的一小部分。

深度优先和广度优先通常是混合使用的，这样既可以照顾到尽量多的网站（广度优先），也能照顾到一部分网站的内页（深度优先）。

3．吸引蜘蛛

由此可见，虽然理论上蜘蛛能爬行和抓取所有页面，但实际上不能、也不会这么做。SEO 人员要想让自己的更多页面被收录，就要想方设法吸引蜘蛛来抓取。既然不能抓取所有页面，蜘蛛所要做的就是尽量抓取重要页面。哪些页面被认为比较重要呢？有几方面影响因素。

网站和页面权重。质量高、资格老的网站被认为权重比较高，这种网站上的页面被爬行的深度也会比较高，所以会有更多内页被收录。

页面更新度。蜘蛛每次爬行都会把页面数据存储起来。如果第二次爬行发现页面与第一次收录的完全一样，说明页面没有更新，蜘蛛也就没有必要经常抓取。如果页面内容经常更新，蜘蛛就会更加频繁地访问这种页面，页面上出现的新链接，也自然会被蜘蛛更快地跟踪，抓取新页面。

导入链接。无论是外部链接还是同一个网站的内部链接，要被蜘蛛抓取，就必须有导入链接进入页面，否则蜘蛛根本没有机会知道页面的存在。高质量的导入链接也经常使页面上的导出链接被爬行深度增加。

与首页点击距离。一般来说网站上权重最高的是首页，大部分外部链接是指向首页的，蜘蛛访问最频繁的也是首页。离首页点击距离越近，页面权重越高，被蜘蛛爬行的机会也越大。

4．地址库

为了避免重复爬行和抓取网址，搜索引擎会建立一个地址库，记录已经被发现还没有抓取的页面，以及已经被抓取的页面。

地址库中的 URL 有几个来源：

（1）人工录入的种子网站。

（2）蜘蛛抓取页面后，从 HTML 中解析出新的链接 URL，与地址库中的数据进行对比，如果是地址库中没有的网址，就存入待访问地址库。

（3）站长通过搜索引擎网页提交表格提交进来的网址。

蜘蛛按重要性从待访问地址库中提取 URL，访问并抓取页面，然后把这个 URL 从待访问地址库中删除，放进已访问地址库中。

大部分主流搜索引擎都提供一个表格，让站长提交网址。不过这些提交来的网址都只是存入地址库而已，是否收录还要看页面重要性如何。搜索引擎所收录的绝大部分页面是蜘蛛自己跟踪链接得到的。可以说提交页面基本上是毫无用处的，搜索引擎更喜欢自己沿着链接发现新页面。

5．文件存储

搜索引擎蜘蛛抓取的数据存入原始页面数据库。其中的页面数据与用户浏览器得到的 HTML 是完全一样的。每个 URL 都有一个独特的文件编号。

6．爬行时的复制内容检测

检测并删除复制内容通常是在下面介绍的预处理过程中进行的，但现在的蜘蛛在爬行和抓取文件时也会进行一定程度的复制内容检测。遇到权重很低的网站上大量转载或抄袭内容时，很可能不再继续爬行。这也就是有的站长在日志文件中发现了蜘蛛，但页面从来没有被真正收录过的原因。

2.4.2 预处理

在一些 SEO 材料中，“预处理”也被简称为“索引”，因为索引是预处理最主要的步骤。

搜索引擎蜘蛛抓取的原始页面，并不能直接用于查询排名处理。搜索引擎数据库中的页面数都在数万亿级别以上，用户输入搜索词后，靠排名程序实时对这么多页面分析相关性，计算量太大，不可能在一两秒内返回排名结果。因此抓取来的页面必须经过预处理，为最后的查询排名做好准备。

和爬行抓取一样，预处理也是在后台提前完成的，用户搜索时感觉不到这个过程。

1．提取文字

现在的搜索引擎还是以文字内容为基础。蜘蛛抓取到的页面中的 HTML 代码，除了用户在浏览器上可以看到的可见文字外，还包含了大量的 HTML 格式标签、JavaScript 程序等无法用于排名的内容。搜索引擎预处理首先要做的就是从 HTML 文件中去除标签、程序，提取出可以用于排名处理的网页面文字内容。

比如下面这段 HTML 代码：

```
<div id="post-1100" class="post-1100 post hentry category-seo">
<div class="posttitle">
<h2><a
href="http://www.chinamyhosting.com/seoblog/2010/04/01/fools -day/"
rel="bookmark" title="Permanent Link to 今天愚人节哈">今天愚人节哈</a></h2>
```

除去 HTML 代码后，剩下的用于排名的文字只是这一行：

```
今天愚人节哈
```

除了可见文字，搜索引擎也会提取出一些特殊的包含文字信息的代码，如 Meta 标签中的文字、图片替代文字、Flash 文件的替代文字、链接锚文字等。

2．中文分词

分词是中文搜索引擎特有的步骤。搜索引擎存储和处理页面及用户搜索都是以词为基础的。英文等语言单词与单词之间有空格分隔，搜索引擎索引程序可以直接把句子划分为单词的集合。而中文词与词之间没有任何分隔符，一个句子中的所有字和词都是连

在一起的。搜索引擎必须首先分辨哪几个字组成一个词，哪些字本身就是一个词。比如“减肥方法”将被分词为“减肥”和“方法”两个词。

中文分词方法基本上有两种，一种是基于词典匹配，另一种是基于统计。

基于词典匹配的方法是指，将待分析的一段汉字与一个事先造好的词典中的词条进行匹配，在待分析汉字串中扫描到词典中已有的词条则匹配成功，或者说切分出一个单词。

按照扫描方向，基于词典的匹配法可以分为正向匹配和逆向匹配。按照匹配长度优先级的不同，又可以分为最大匹配和最小匹配。将扫描方向和长度优先混合，又可以产生正向最大匹配、逆向最大匹配等不同方法。

词典匹配方法计算简单，其准确度在很大程度上取决于词典的完整性和更新情况。

基于统计的分词方法指的是分析大量文字样本，计算出字与字相邻出现的统计概率，几个字相邻出现越多，就越可能形成一个单词。基于统计的方法的优势是对新出现的词反应更快速，也有利于消除歧义。

基于词典匹配和基于统计的分词方法各有优劣，实际使用中的分词系统都是混合使用两种方法的，快速高效，又能识别生词、新词，消除歧义。

中文分词的准确性往往影响搜索引擎排名的相关性。比如在百度搜索“搜索引擎优化”，如图 2-22 所示，从快照中可以看到，百度把“搜索引擎优化”这六个字当成一个词。

图 2-22　百度快照显示的对“搜索引擎优化”的分词结果

而在 Google 搜索同样的词，如图 2-23 所示，快照显示 Google 将其分切为“搜索”、“引擎”和“优化”三个词。显然百度切分得更为合理，搜索引擎优化是一个完整的概念。Google 分词时倾向于更为细碎。

图 2-23　Google 快照显示的对“搜索引擎优化”的分词结果

再举一个更明显的例子。在 Google 搜索“点石互动”四个字，如图 2-24 所示，快照显示 Google 将其切分为“点”，“石”及“互动”三个词。“点石互动”这个中文 SEO 领域最知名的品牌，显然并没有进入 Google 的词典中。

图 2-24　Google 快照显示的对“点石互动”的分词结果

在百度搜索“点石互动”时会发现，百度将“点石互动”当做一个词。甚至在百度搜索“点石大会报名”，可以发现百度把“点石大会”都当成一个词，如图 2-25 所示。

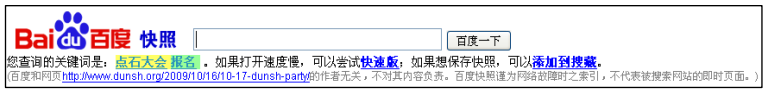


图 2-25　百度快照显示的对“点石大会报名”的分词结果

这种分词上的不同很可能是一些关键词排名在不同搜索引擎有不同表现的原因之一。比如百度更喜欢将搜索词完整匹配地出现在页面上，也就是说搜索“点石互动”时，这四个字连续完整出现更容易在百度获得好的排名。Google 就与此不同，不太要求完整匹配。一些页面出现“点石”和“互动”两个词，但不必完整匹配地出现，“点石”出现在前面，“互动”出现在页面的其他地方，这样的页面在 Google 搜索“点石互动”时，也可以获得不错的排名。

需要说明的是，从快照中只能在某些情况下大致猜测搜索引擎的分词处理，很多时候我们并不能从快照中看出搜索引擎是怎么分词的。比如在百度中搜索一些很长的词甚至句子，对于不同网站，百度快照有时候将整个句子标为同一个颜色，有时候又分为不同颜色的词的组合。即使在快照中将整个句子标为同一个颜色，显然百度也不是将整个句子分为一个词。

搜索引擎对页面的分词取决于词库的规模、准确性和分词算法的好坏，而不是取决于页面本身如何，所以 SEO 人员对分词所能做的很少。唯一能做的是在页面上用某种形式提示搜索引擎，某几个字应该被当做一个词处理，尤其是可能产生歧义的时候，比如在页面标题、h1 标签及黑体中出现关键词。如果页面是关于“和服”的内容，那么可以把“和服”这两个字特意标为黑体。如果页面是关于“化妆和服装”，可以把“服装”两个字标为黑体。这样，搜索引擎对页面进行分析时就知道标为黑体的应该是一个词。

3．去停止词

无论是英文还是中文，页面内容中都会有一些出现频率很高，却对内容没有任何影响的词，如“的”、“地”、“得”之类的助词，“啊”、“哈”、“呀”之类的感叹词，“从而”、“以”、“却”之类的副词或介词。这些词被称为停止词，因为它们对页面的主要意思没什么影响。英文中的常见停止词有 the，a，an，to，of 等。

搜索引擎在索引页面之前会去掉这些停止词，使索引数据主题更为突出，减少无谓的计算量。

4．消除噪声

绝大部分页面上还有一部分内容对页面主题也没有什么贡献，比如版权声明文字、导航条、广告等。以常见的博客导航为例，几乎每个博客页面上都会出现文章分类、历史存档等导航内容，但是这些页面本身与“分类”、“历史”这些词都没有任何关系。用

户搜索“历史”、“分类”这些关键词时仅仅因为页面上有这些词出现而返回博客帖子是毫无意义的，完全不相关。所以这些区块都属于噪声，对页面主题只能起到分散作用。

搜索引擎需要识别并消除这些噪声，排名时不使用噪声内容。消噪的基本方法是根据 HTML 标签对页面分块，区分出页头、导航、正文、页脚、广告等区域，在网站上大量重复出现的区块往往属于噪声。对页面进行消噪后，剩下的才是页面主体内容。

5．去重

搜索引擎还需要对页面进行去重处理。

同一篇文章经常会重复出现在不同网站及同一个网站的不同网址上，搜索引擎并不喜欢这种重复性的内容。用户搜索时，如果在前两页看到的都是来自不同网站的同一篇文章，用户体验就太差了，虽然都是内容相关的。搜索引擎希望只返回相同文章中的一篇，所以在进行索引前还需要识别和删除重复内容，这个过程就称为“去重”。

“去重”的基本方法是对页面特征关键词计算指纹，也就是说从页面主体内容中选取最有代表性的一部分关键词（经常是出现频率最高的关键词），然后计算这些关键词的数字指纹。这里的关键词选取是在分词、去停止词、消噪之后。实验表明，通常选取 10 个特征关键词就可以达到比较高的计算准确性，再选取更多词对去重准确性提高的贡献也就不大了。

典型的指纹计算方法如 MD5 算法（信息摘要算法第五版）。这类指纹算法的特点是，输入（特征关键词）有任何微小的变化，都会导致计算出的指纹有很大差距。

了解了搜索引擎的去重算法，SEO 人员就应该知道简单地增加“的”、“地”、“得”、调换段落顺序这种所谓伪原创，并不能逃过搜索引擎的去重算法，因为这样的操作无法改变文章的特征关键词。而且搜索引擎的去重算法很可能不止于页面级别，而是进行到段落级别，混合不同文章、交叉调换段落顺序也不能使转载和抄袭变成原创。

6．正向索引

正向索引也可以简称为索引。

经过文字提取、分词、消噪、去重后，搜索引擎得到的就是独特的、能反映页面主体内容的、以词为单位的内容。接下来搜索引擎索引程序就可以提取关键词，按照分词程序划分好的词，把页面转换为一个关键词组成的集合，同时记录每一个关键词在页面上的出现频率、出现次数、格式（如出现在标题标签、黑体、H 标签、锚文字等）、位置（如页面第一段文字等）。这样，每一个页面都可以记录为一串关键词集合，其中每个关键词的词频、格式、位置等权重信息也都记录在案。

搜索引擎索引程序将页面及关键词形成词表结构存储进索引库。简化的索引词表形式如表 2-1 所示。

表 2-1　简化的索引词表结构

文件 ID	内　　容
文件 1	关键词 1，关键词 2，关键词 7，关键词 10，……，关键词 *L*
文件 2	关键词 1，关键词 7，关键词 30，……，关键词 *M*
文件 3	关键词 2，关键词 70，关键词 305，……，关键词 *N*
⋮	
文件 6	关键词 2，关键词 7，关键词 10，……，关键词 *X*
⋮	
文件 *x*	关键词 7，关键词 50，关键词 90，……，关键词 *Y*

每个文件都对应一个文件 ID，文件内容被表示为一串关键词的集合。实际上在搜索引擎索引库中，关键词也已经转换为关键词 ID。这样的数据结构就称为正向索引。

7．倒排索引

正向索引还不能直接用于排名。假设用户搜索关键词 2，如果只存在正向索引，排名程序需要扫描所有索引库中的文件，找出包含关键词 2 的文件，再进行相关性计算。这样的计算量无法满足实时返回排名结果的要求。

所以搜索引擎会将正向索引数据库重新构造为倒排索引，把文件对应到关键词的映射转换为关键词到文件的映射，如表 2-2 所示。

在倒排索引中关键词是主键，每个关键词都对应着一系列文件，这些文件中都出现了这个关键词。这样当用户搜索某个关键词时，排序程序在倒排索引中定位到这个关键词，就可以马上找出所有包含这个关键词的文件。

表 2-2　倒排索引结构

关 键 词	文　　件
关键词 1	文件 1，文件 2，文件 15，文件 58，……，文件 *l*
关键词 2	文件 1，文件 3，文件 6，……，文件 *m*
关键词 3	文件 5，文件 700，文件 805，……，文件 *n*
⋮	
关键词 7	文件 1，文件 2，文件 6，……，文件 *x*
⋮	
关键词 *Y*	文件 80，文件 90，文件 100，……，文件 *x*

8．链接关系计算

链接关系计算也是预处理中很重要的一部分。现在所有的主流搜索引擎排名因素中都包含网页之间的链接流动信息。搜索引擎在抓取页面内容后，必须事前计算出：页面

上有哪些链接指向哪些其他页面，每个页面有哪些导入链接，链接使用了什么锚文字，这些复杂的链接指向关系形成了网站和页面的链接权重。

Google PR 值就是这种链接关系的最主要体现之一。其他搜索引擎也都进行类似计算，虽然它们并不称为 PR。

由于页面和链接数量巨大，网上的链接关系又时时处在更新中，因此链接关系及 PR 的计算要耗费很长时间。关于 PR 和链接分析，后面还有专门的章节介绍。

9．特殊文件处理

除了 HTML 文件外，搜索引擎通常还能抓取和索引以文字为基础的多种文件类型，如 PDF、Word、WPS、XLS、PPT、TXT 文件等。我们在搜索结果中也经常会看到这些文件类型。但目前的搜索引擎还不能处理图片、视频、Flash 这类非文字内容，也不能执行脚本和程序。

虽然搜索引擎在识别图片及从 Flash 中提取文字内容方面有些进步，不过距离直接靠读取图片、视频、Flash 内容返回结果的目标还很远。对图片、视频内容的排名还往往是依据与之相关的文字内容，详细情况可以参考后面的整合搜索部分。

2.4.3 排名

经过搜索引擎蜘蛛抓取页面，索引程序计算得到倒排索引后，搜索引擎就准备好可以随时处理用户搜索了。用户在搜索框填入关键词后，排名程序调用索引库数据，计算排名显示给用户，排名过程是与用户直接互动的。

1．搜索词处理

搜索引擎接收到用户输入的搜索词后，需要对搜索词做一些处理，才能进入排名过程。搜索词处理包括如下几方面。

（1）中文分词。与页面索引时一样，搜索词也必须进行中文分词，将查询字符串转换为以词为基础的关键词组合。分词原理与页面分词相同。

（2）去停止词。和索引时一样，搜索引擎也需要把搜索词中的停止词去掉，最大限度地提高排名相关性及效率。

（3）指令处理。查询词完成分词后，搜索引擎的默认处理方式是在关键词之间使用“与”逻辑。也就是说用户搜索“减肥方法”时，程序分词为“减肥”和“方法”两个词，搜索引擎排序时默认认为，用户寻找的是既包含“减肥”，也包含“方法”的页面。只包含“减肥”不包含“方法”，或者只包含“方法”不包含“减肥”的页面，被认为是不符合搜索条件的。当然，这只是极为简化的为了说明原理的说法，实际上我们还是会看到只包含一部分关键词的搜索结果。

另外用户输入的查询词还可能包含一些高级搜索指令，如加号、减号等，搜索引擎都需要做出识别和相应处理。有关高级搜索指令，后面还有详细说明。

（4）拼写错误矫正。用户如果输入了明显错误的字或英文单词拼错，搜索引擎会提示用户正确的用字或拼法，如图 2-26 所示。

图 2-26　输入的错拼、错字矫正

（5）整合搜索触发。某些搜索词会触发整合搜索，比如明星姓名就经常触发图片和视频内容，当前的热门话题又容易触发资讯内容。哪些词触发哪些整合搜索，也需要在搜索词处理阶段计算。

2．文件匹配

搜索词经过处理后，搜索引擎得到的是以词为基础的关键词集合。文件匹配阶段就是找出含有所有关键词的文件。在索引部分提到的倒排索引使得文件匹配能够快速完成，如表 2-3 所示。

表 2-3　倒排索引快速匹配文件

关　键　词	文　　件
关键词 1	文件 1，文件 2，文件 15，文件 58，……，文件 *l*
关键词 2	**文件 1**，文件 3，**文件 6**，……，文件 *m*
关键词 3	文件 5，文件 700，文件 805，……，文件 *n*
⋮	
关键词 7	**文件 1**，文件 2，**文件 6**，……，文件 *x*
⋮	
关键词 *Y*	文件 80，文件 90，文件 100，……，文件 *x*

假设用户搜索“关键词 2”和“关键词 7”，排名程序只要在倒排索引中找到“关键词 2”和“关键词 7”这两个词，就能找到分别含有这两个词的所有页面。经过简单计算就能找出既包含“关键词 2”，也包含“关键词 7”的所有页面：文件 1 和文件 6。

3．初始子集的选择

找到包含所有关键词的匹配文件后，还不能进行相关性计算，因为找到的文件经常

会有几十万几百万，甚至上千万个。要对这么多文件实时进行相关性计算，需要的时间还是比较长的。

实际上用户并不需要知道所有匹配的几十万、几百万个页面，绝大部分用户只会查看前两页，也就是前 20 个结果。搜索引擎也并不需要计算这么多页面的相关性，而只要计算最重要的一部分页面就可以了。常用搜索引擎的人都会注意到，搜索结果页面通常最多显示 100 个。用户点击搜索结果页面底部的“下一页”链接，最多也只能看到第 100 页，也就是 1000 个搜索结果，如图 2-27 所示。

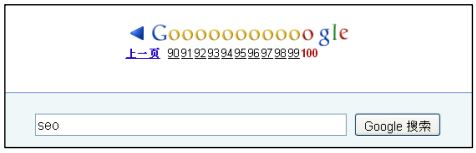


图 2-27　Google 显示 100 页搜索结果

百度则通常返回 76 页结果，如图 2-28 所示。

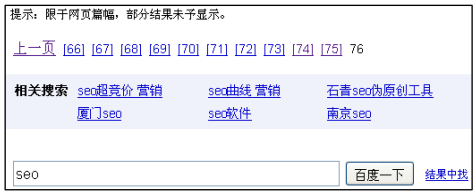


图 2-28　百度显示 76 页搜索结果

所以搜索引擎只需要计算前 1000 个结果的相关性，就能满足要求。

但问题在于，还没有计算相关性时，搜索引擎又怎么知道哪 1000 个文件是最相关的？所以用于最后相关性计算的初始页面子集的选择，必须依靠其他特征而不是相关性，其中最主要的就是页面权重。由于所有匹配文件都已经具备了最基本的相关性（这些文件都包含所有查询关键词），搜索引擎通常会用非相关性的页面特征选出一个初始子集。初始子集的数目是多少？几万个？或许更多，外人并不知道。不过可以肯定的是，当匹配页面数目巨大时，搜索引擎不会对这么多页面进行计算，而必须选出页面权重较高的一个子集，再对子集中的页面进行相关性计算。

4．相关性计算

选出初始子集后，对子集中的页面计算关键词相关性。计算相关性是排名过程中最重要的一步。相关性计算是搜索引擎算法中最令 SEO 感兴趣的部分。

影响相关性的主要因素包括如下几方面。

（1）关键词常用程度。经过分词后的多个关键词，对整个搜索字符串的意义贡献并不相同。越常用的词对搜索词的意义贡献越小，越不常用的词对搜索词的意义贡献越大。

举个例子，假设用户输入的搜索词是“我们冥王星”。“我们”这个词常用程度非常高，在很多页面上会出现，它对“我们冥王星”这个搜索词的辨识程度和意义相关度贡献就很小。找出那些包含“我们”这个词的页面，对搜索排名相关性几乎没有什么影响，有太多页面包含“我们”这个词。

而“冥王星”这个词常用程度就比较低，对“我们冥王星”这个搜索词的意义贡献要大得多。那些包含“冥王星”这个词的页面，对“我们冥王星”这个搜索词会更为相关。

常用词的极致就是停止词，对页面意义完全没有影响。

所以搜索引擎对搜索词串中的关键词并不是一视同仁地处理，而是根据常用程度进行加权。不常用的词加权系数高，常用词加权系数低，排名算法对不常用的词给予更多关注。

我们假设 A、B 两个页面都各出现“我们”及“冥王星”两个词。但是“我们”这个词在 A 页面出现于普通文字中，“冥王星”这个词在 A 页面出现于标题标签中。B 页面正相反，“我们”出现在标题标签中，而“冥王星”出现在普通文字中。那么针对“我们冥王星”这个搜索词，A 页面将更相关。

（2）词频及密度。一般认为在没有关键词堆积的情况下，搜索词在页面中出现的次数多，密度越高，说明页面与搜索词越相关。当然这只是一个大致规律，实际情况未必如此，所以相关性计算还有其他因素。出现频率及密度只是因素的一部分，而且重要程度越来越低。

（3）关键词位置及形式。就像在索引部分中提到的，页面关键词出现的格式和位置都被记录在索引库中。关键词出现在比较重要的位置，如标题标签、黑体、H1 等，说明页面与关键词越相关。这一部分就是页面 SEO 所要解决的。

（4）关键词距离。切分后的关键词完整匹配地出现，说明与搜索词最相关。比如搜索“减肥方法”时，页面上连续完整出现“减肥方法”四个字是最相关的。如果“减肥”和“方法”两个词没有连续匹配出现，出现的距离近一些，也被搜索引擎认为相关性稍微大一些。

（5）链接分析及页面权重。除了页面本身的因素，页面之间的链接和权重关系也影响关键词的相关性，其中最重要的是锚文字。页面有越多以搜索词为锚文字的导入链接，说明页面的相关性越强。

链接分析还包括了链接源页面本身的主题、锚文字周围的文字等。

上面简单介绍的几个因素在本书中都有更详细的说明。

5．排名过滤及调整

选出匹配文件子集、计算相关性后，大体排名就已经确定了。之后搜索引擎可能还

有一些过滤算法，对排名进行轻微调整，其中最主要的过滤就是施加惩罚。一些有作弊嫌疑的页面，虽然按照正常的权重和相关性计算排到前面，但搜索引擎的惩罚算法却可能在最后一步把这些页面调到后面去。典型的例子是百度的 11 位，Google 的负 6、负 30、负 950 等算法。

6．排名显示

所有排名确定后，排名程序调用原始页面的标题标签、说明标签、快照日期等数据显示在页面上。有时搜索引擎需要动态生成页面摘要，而不是调用页面本身的说明标签。

7．搜索缓存

用户搜索的关键词有很大一部分是重复的。按照 2/8 定律，20%的搜索词占到了总搜索次数的 80%。按照长尾理论，最常见的搜索词没有占到 80%那么多，但通常也有一个比较粗大的头部，很少一部分搜索词占到了所有搜索次数的很大一部分。尤其是有热门新闻发生时，每天可能有几百万人搜索完全相同的关键词。

如果每次搜索都重新处理排名可以说是很大的浪费。搜索引擎会把最常见的搜索词存入缓存，用户搜索时直接从缓存中调用，而不必经过文件匹配和相关性计算，大大提高了排名效率，缩短了搜索反应时间。

8．查询及点击日志

搜索用户的 IP 地址、搜索的关键词、搜索时间，以及点击了哪些结果页面，搜索引擎都记录形成日志。这些日志文件中的数据对搜索引擎判断搜索结果质量、调整搜索算法、预期搜索趋势等都有重要意义。

上面我们简单介绍了搜索引擎的工作过程。当然实际搜索引擎的工作步骤与算法是非常复杂的。上面的说明很简单，但其中有很多技术难点。

搜索引擎还在不断优化算法，优化数据库格式。不同搜索引擎的工作步骤也会有差异。但大致上所有主流搜索引擎的基本工作原理都是如此，在过去几年及可以预期的未来几年，都不会有实质性的改变。

2.5 链接原理

在 Google 诞生以前，传统搜索引擎主要依靠页面内容中的关键词匹配搜索词进行排名。这种排名方式的短处现在看来显而易见，那就是很容易被刻意操纵。黑帽 SEO 在页面上堆积关键词，或加入与主题无关的热门关键词，都能提高排名，使搜索引擎排名结果质量大为下降。现在的搜索引擎都使用链接分析技术减少垃圾，提高用户体验。本节就简要探讨链接在搜索引擎排名中的应用原理。

在排名中计入链接因素，不仅有助于减少垃圾，提高结果相关性，也使传统关键词匹配无法排名的文件能够被处理。比如图片、视频文件无法进行关键词匹配，但是却可能有外部链接，通过链接信息，搜索引擎就可以了解图片和视频的内容从而排名。

不同文字的页面排名也成为可能。比如在百度或 google.cn 搜索“SEO”，都可以看到英文和其他文字的 SEO 网站。甚至搜索“搜索引擎优化”，也可以看到非中文页面，原因就在于有的链接可能使用“搜索引擎优化”为锚文字指向英文页面。

链接因素现在已经超过页面内容的重要性。不过理解链接关系比较抽象。页面上的因素对排名的影响能看得到，容易直观理解。举个简单的例子，搜索一个特定关键词，SEO 人员只要观察前几页结果，就能看到：关键词在标题标签中出现有什么影响，出现在最前面又有什么影响，有技术资源的还可以大规模地统计，计算出关键词出现在标题标签中不同位置与排名之间的关系。虽然这种关系不一定是因果关系，但至少是统计上的联系，使 SEO 人员大致了解如何优化。

链接对排名的影响就无法直观了解，也很难进行统计，因为没有人能获得搜索引擎的链接数据库。我们能做的最多只是定性观察和分析。

下面介绍的一些关于链接的专利，多少透露了链接在搜索引擎排名中的使用方法和地位。

2.5.1 李彦宏超链分析专利

百度创始人李彦宏在回国创建百度之前就是美国顶级的搜索引擎工程师之一。据说李彦宏在寻找风险投资时，投资人询问其他三个搜索引擎业界的技术高人一个问题：要了解搜索引擎技术应该问谁。这三个被问到的高人中有两个回答：搜索引擎的事就问李彦宏。由此投资人断定李彦宏是最了解搜索引擎的人之一。

这其实就是现实生活中类似于链接关系的应用。要判断哪个页面最具权威性，不能光看页面自己怎么说，还要看其他页面怎么评价。

李彦宏 1997 年就提交了一份名为“超链文件检索系统和方法”的专利申请，这比 Google 创始人发明 PR 要早得多，不得不说这是非常具有前瞻性的研究工作。在这份专利中，李彦宏提出了与传统信息检索系统不同的基于链接的排名方法。

这个系统除了索引页面之外，还建立一个链接词库，记录链接锚文字的一些相关信息，如锚文字中包含哪些关键词，发出链接的页面索引，包含特定锚文字的链接总数，包含特定关键词的链接都指向哪些页面。词库不仅包含关键词原型，也包含同一个词干的其他衍生关键词。

根据这些链接数据，尤其是锚文字，计算出基于链接的文件相关性。在用户搜索时，将得到的基于链接的相关性与基于关键词匹配的传统相关性综合使用，得到更准确的排名。

在今天看来，这种基于链接的相关性计算是搜索引擎的常态，每个 SEO 人员都知道。但是在十三四年前，这无疑是非常创新的概念。当然现在的搜索引擎算法对链接的考虑，已经不仅仅是锚文字，实际上要复杂得多。

这份专利所有人是李彦宏当时所在的公司，发明人是李彦宏本人。感兴趣的读者可以在下面这个地址查看美国专利局发布的“超链文件检索系统和方法”专利详情：

http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=5,920,859

2.5.2 HITS 算法

HITS 是英文 Hyperlink-Induced Topic Search 的缩写，意译为“超链诱导主题搜索”。HITS 算法由 Jon Kleinberg 于 1997 年提出，并申请了专利：

http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=6,112,202

按照 HITS 算法，用户输入关键词后，算法对返回的匹配页面计算两种值，一种是枢纽值（Hub Scores），另一种是权威值（Authority Scores），这两个值是互相依存、互相影响的。所谓枢纽值，指的是页面上所有导出链接指向页面的权威值之和。权威值指的是所有导入链接所在页面的枢纽值之和。

上面的定义比较拗口，我们可以简单地说，HITS 算法会提炼出两种比较重要的页面，也就是枢纽页面和权威页面。枢纽页面本身可能没有多少导入链接，但是有很多导出链接指向权威页面。权威页面本身可能导出链接不多，但是有很多来自枢纽页面的导入链接。

典型的枢纽页面就是如雅虎目录、开放目录或好 123 这样的网站目录。这种高质量的网站目录作用就在于指向其他权威网站，所以称为枢纽。而权威页面有很多导入链接，其中包含很多来自枢纽页面的链接。权威页面通常是提供真正相关内容的页面。

HITS 算法是针对特定查询词的，所以称为主题搜索。

HITS 算法的最大缺点是，它在查询阶段进行计算，而不是在抓取或预处理阶段。所以 HITS 算法是以牺牲查询排名响应时间为代价的。也正因为如此，原始 HITS 算法在搜索引擎中并不常用。不过 HITS 算法的思想很可能融入到搜索引擎的索引阶段，也就是根据链接关系找出具有枢纽特征或权威特征的页面。

成为权威页面是第一优先，不过难度比较大，唯一的方法就是获得高质量链接。当你的网站不能成为权威页面时，就让它成为枢纽页面。所以导出链接也是当前搜索引擎排名因素之一。绝不链接到其他网站的做法，并不是好的 SEO 方法。

2.5.3 TrustRank 算法

TrustRank 是近年来比较受关注的基于链接关系的排名算法。TrustRank 可以翻译为“信任指数”。

TrustRank 算法最初来自于 2004 年斯坦福大学和雅虎的一项联合研究，用来检测垃圾网站，并且于 2006 年申请专利。TrustRank 算法发明人还发表了一份专门的 PDF 文件，说明 TrustRank 算法的应用。感兴趣的读者可以在下面这个网址下载 PDF 文件：

http://www.vldb.org/conf/2004/RS15P3.PDF

TrustRank 算法并不是由 Google 提出的，不过由于 Google 所占市场份额最大，而且 TrustRank 在 Google 排名中也是一个非常重要的因素，所以有些人误以为 TrustRank 是 Google 提出的。更让人糊涂的是，Google 曾经把 TrustRank 申请为商标，但是 TrustRank 商标中的 TrustRank 指的是 Google 检测含有恶意代码网站的方法，而不是指排名算法中的信任指数。

TrustRank 算法基于一个基本假设：好的网站很少会链接到坏的网站。反之则不成立，也就是说，坏的网站很少链接到好网站这句话并不成立。正相反，很多垃圾网站会链接到高权威、高信任指数的网站，试图提高自己的信任指数。

基于这个假设，如果能挑选出可以百分之百信任的网站，这些网站的 TrustRank 评为最高，这些 TrustRank 最高的网站所链接到的网站信任指数稍微降低，但也会很高。与此类似，第二层被信任的网站链接出去的第三层网站，信任度继续下降。由于种种原因，好的网站也不可避免地会链接到一些垃圾网站，不过离第一层网站点击距离越近，所传递的信任指数越高，离第一级网站点击距离就越远，信任指数将依次下降。这样，通过 TrustRank 算法，就能给所有网站计算出相应的信任指数，离第一层网站越远，成为垃圾网站的可能性就越大。

计算 TrustRank 值首先要选择一批种子网站，然后人工查看网站，设定一个初始 TrustRank 值。挑选种子网站有两种方式，一种是选择导出链接最多的网站，因为 TrustRank 算法就是计算指数随着导出链接的衰减。导出链接多的网站，在某种意义上可以理解为“逆向 PR 值”比较高。

另一种挑选种子网站的方法是选 PR 值高的网站，因为 PR 值越高，在搜索结果页面出现的概率就越大。这些网站才正是 TrustRank 算法最关注的、需要调整排名的网站。那些 PR 值很低的页面，在没有 TrustRank 算法时排名也很靠后，计算 TrustRank 意义就不大了。

根据测算，挑选出两百个左右网站作为种子，就可以比较精确地计算出所有网站的 TrustRank 值。

计算 TrustRank 随链接关系减少的公式有两种方式。一种是随链接次数衰减，也就是说如果第一层页面 TrustRank 指数是 100，第二层页面衰减为 90，第三层衰减为 80。第二种计算方法是按导出链接数目分配 TrustRank 值，也就是说，如果一个页面的 TrustRank 值是 100，页面上有 5 个导出链接，每个链接将传递 20%的 TrustRank 值。衰减和分配这两种计算方法通常综合使用，整体效果都是随着链接层次的增加，TrustRank 值逐步降低。

得出网站和页面的 TrustRank 值后，可以通过两种方式影响排名。一种是把传统排名算法挑选出的多个页面，根据 TrustRank 值比较，重新做排名调整。另一种是设定一个最低的 TrustRank 值门槛，只有超过这个门槛的页面，才被认为有足够的质量进入排名，低于门槛的页面将被认为是垃圾页面，从搜索结果中过滤出去。

虽然 TrustRank 算法最初是作为检测垃圾的方法，但在现在的搜索引擎排名算法中，TrustRank 概念使用更为广泛，常常影响大部分网站的整体排名。TrustRank 算法最初针对的是页面级别，现在在搜索引擎算法中，TrustRank 值也通常表现在域名级别，整个域名的信任指数越高，整体排名能力就越强。

2.5.4 Google PR

PR 是 PageRank 的缩写。Google PR 理论是所有基于链接的搜索引擎理论中最有名的。SEO 人员可能不清楚本节介绍的其他链接理论，但不可能不知道 PR。

PR 是 Google 创始人之一拉里佩奇发明的，用于表示页面重要性的概念。用最简单的话说就是，反向链接越多的页面就是越重要的页面，因此 PR 值也就越高。

Google PR 有点类似于科技文献中互相引用的概念，被其他文献引用较多的文献，很可能是比较重要的文献。

1．PR 的概念和计算

我们可以把互联网理解为由结点及链接组成的有向图，页面就是一个个结点，页面之间的有向链接传递着页面的重要性。一个链接传递的 PR 值决定于导入链接所在页面的 PR 值，发出链接的页面本身 PR 值越高，所能传递出去的 PR 也越高。传递的 PR 数值也取决于页面上的导出链接数目。对于给定 PR 值的页面来说，假设能传递到下级页面 100 份 PR，页面上有 10 个导出链接，每个链接能传递 10 份 PR，页面上有 20 个导出链接的话，每个链接只能传递 5 份 PR。所以一个页面的 PR 值取决于导入链接总数，发出链接页面的 PR 值，以及发出链接页面上的导出链接数目。

PR 值计算公式是：

$$\mathrm{PR}(A) = (1-d) + d(\mathrm{PR}(t_1)/C(t_1) + \cdots + \mathrm{PR}(t_n)/C(t_n))$$

A 代表页面 A。

PR(A)则代表页面 A 的 PR 值。

d 为阻尼指数。通常认为 d=0.85。

$t_1 \cdots t_n$ 代表链接向页面 A 的页面 t_1 到 t_n。

C 代表页面上的导出链接数目。$C(t_1)$即为页面 t_1 上的导出链接数目。

从概念及计算公式都可以看到，计算 PR 值必须使用迭代计算。页面 A 的 PR 值取决于链接向 A 的页面 t_1 至 t_n 页面的 PR 值，而 t_1 至 t_n 页面的 PR 值又取决于其他页面的 PR 值，其中很可能还包含页面 A。所以 PR 需要多次迭代才能得到。计算时先给所有页面设定一个初始值，经过一定次数的迭代计算后，各个页面的 PR 值将趋于稳定。研究证明，无论初始值怎么选取，经过迭代计算的最终 PR 值不会受到影响。

对阻尼系数做个简要说明。考虑如图 2-29 所示这样一个循环（实际网络上是一定存在这种循环的）。

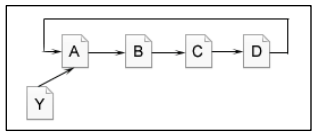


图 2-29　链接构成的循环

外部页面 Y 向循环注入 PR 值，循环中的页面不停迭代传递 PR，如果没有阻尼系数，循环中的页面 PR 将达到无穷大。引入阻尼系数，使 PR 在传递时自然衰减，才能将 PR 计算稳定在一个值上。

2．PR 的两个比喻模型

关于 PR 有两个著名的比喻。一个比喻是投票。链接就像民主投票一样，A 页面链接到 B 页面，就意味着 A 页面对 B 页面投了一票，使得 B 页面的重要性提高。同时，A 页面本身的 PR 值决定了 A 所能投出去的投票力，PR 值越高的页面，投出的票也更重要。在这个意义上，传统基于关键词匹配的算法是看页面自己说页面内容是什么，基于链接的 PR 则是看别人怎么评价一个页面。

第二个比喻是随机冲浪比喻。假设一个访问者从一个页面开始，不停地随机点击链接，访问下一个页面。有时候这个用户感到无聊了，不再点击链接，就随机跳到了另外一个网址，再次开始不停地向下点击。所谓 PR 值，也就是一个页面在这种随机冲浪访问中被访问到的概率。一个页面导入链接越多，被访问到的概率也越高，因此 PR 值也越高。

阻尼系数也与随机冲浪模型有关。(1–d)=0.15 实际上就是用户感到无聊，停止点击，

随机跳到新 URL 的概率。

3．工具条 PR

真正的用于排名计算的 Google PR 值我们是无法知道的，我们所能看到的只是 Google 工具条 PR 值。需要清楚的是，工具条 PR 值并不是真实 PR 值的精确反映。真实 PR 值是一个准确的、大于 0.15、没有上限的数字，工具条上显示的 PR 值已经简化为 0～10 这 11 个数字，是一个整数，也就是说 PR 值最小的近似为 0，最大的近似为 10。实际上每一个工具条 PR 值代表的是很大一个范围，工具条 PR5 代表的页面真实 PR 值可能相差很多倍。

真正的 PR 值是不间断计算更新中的，工具条 PR 值只是某一个时间点上真实 PR 值的快照输出。工具条 PR 几个月才更新一次，过去一年工具条 PR 值更新的日期如表 2-4 所示。

表 2-4　最近一年 Google 工具条 PR 值更新日期

2010 年 4 月 1 号
2009 年 12 月 31 号
2009 年 10 月 29 号
2009 年 6 月 23 号
2009 年 5 月 26 号
2009 年 4 月 1 号
2008 年 12 月 31 号

工具条 PR 与反向链接数目呈对数关系，而不是线性关系。也就是说如果从 PR1 到 PR2 需要的外部链接是 100 个，从 PR2 到 PR3 则需要大致 1000 个，PR5 到 PR6 需要的外部链接则更多。所以 PR 值越高的网站想提升一级所要付出的时间和努力比 PR 值比较低的网站提升一级要多得多。

4．关于 PR 的几个误解

PR 的英文全称是 PageRank。这个名称来源于发明人佩奇（Page）的名字，巧合的是 Page 在英文中也是页面的意思。所以准确地说 PageRank 这个名字应该翻译为佩奇级别，而不是页面级别。不过约定俗成，再加上巧妙的一语双关，大家都把 PR 称为页面级别。

PR 值只与链接有关。经常有站长询问，他的网站做了挺长时间，内容也全是原创的，怎么 PR 还是零呢？其实 PR 与站长是否认真、做站多长时间、内容是否原创都没有直接关系。有反向链接就有 PR，没有反向链接就没有 PR。一个高质量的原创网站，一般来说自然会吸引到比较多的外部链接，所以会间接提高 PR 值，但这并不是必然的。

工具条 PR 值更新与页面排名变化在时间上没有对应关系。在工具条 PR 值更新过程中，经常有站长说 PR 值提高了，难怪网站排名也提高了。可以肯定地说这只是时间上的巧合而已。前面说过，真实的用于排名计算的 PR 是连续计算更新的，随时计入排名算法。我们看到的工具条 PR 几个月才更新一次。当我们看到有 PR 更新时，真实的 PR 早在几个月之前就更新和计入排名里了。所以，通过工具条 PR 变化，研究 PR 值与排名变化之间的关系是没有意义的。

5．PR 的意义

Google 工程师说过很多次，Google PR 现在已经是一个被过度宣传的概念，其实 PR 只是 Google 排名算法的 200 多个因素之一，而且重要性已经下降很多，SEO 人员完全不必太执着于 PR 值的提高。

当然，PR 还是 Google 排名算法中的重要因素之一。除了直接影响排名，PR 的重要性还体现在下面几点。

（1）网站收录深度和总页面数。搜索引擎蜘蛛爬行时间及数据库的空间都是有限的。Google 希望尽量优先收录重要性高的页面，所以 PR 值越高的网站就能被收录更多页面，蜘蛛爬行内页的深度也更高。对大中型网站来说，首页 PR 值是带动网站收录的重要因素之一。

（2）更新频率。PR 值越高的网站，搜索引擎蜘蛛访问得就越频繁，网站上出现新页面或旧页面上内容更新时，都能更快速地被收录。由于网站新页面通常都会在现有页面上出现链接，因此更新频率高也就意味着被发现的速度快。

（3）重复内容判定。当 Google 在不同网站上发现完全相同的内容时，会选出一个作为原创，其他作为转载或抄袭。用户搜索相关关键词时，被判断为原创的那个版本会排在前面。而判断哪个版本为原创时，PR 值也是重要因素之一。这也就是为什么那些权重高、PR 值高的大网站，转载小网站内容却经常被当做原创的原因。

（4）排名初始子集的选择。前面介绍排名过程时提到，搜索引擎挑选出所有与关键词匹配的文件后，不可能对所有文件进行相关性计算，因为返回的文件可能有几百万、几千万，搜索引擎需要从中挑选出一个初始子集，再做相关性计算。初始子集的选择显然与关键词相关度无关，而只能从页面的重要程度着手，PR 值就是与关键词无关的重要度指标。

现在的 PR 算法与当初拉里佩奇专利中的描述相比肯定有了改进和变化。一个可以观察到的现象是，PR 算法应该已经排除了一部分 Google 认为可疑或者无效的链接，比如付费链接、博客和论坛中的垃圾链接等。所以有时候我们会看到一个页面有 PR6 甚至 PR7 的导入链接，经过几次工具条 PR 更新后，却还维持在 PR3 甚至 PR2。按说一个 PR6 或 7 的链接，应该把被链接的页面带到 PR5 或 PR4，所以很可能 Google 已经把一部分它认为可疑的链接排除在 PR 计算之外。

PR 专利发明人是拉里佩奇，专利所有人是斯坦福大学，Google 公司拥有永久性排他使用权。虽然 PR 是 Google 拥有专利使用权的算法，但其他所有主流搜索引擎也都有类似算法，只不过不称为 PR 而已。

2.5.5 Hilltop 算法

Hilltop 算法是由 Krishna Baharat 在 2000 年左右所研究的，于 2001 年申请了专利，并且把专利授权给 Google 使用，后来 Krishna Baharat 本人也加入了 Google。

Hilltop 算法可以简单理解为与主题相关的 PR 值。传统 PR 值与特定关键词或主题没有关联，只计算链接关系。这就有可能出现某种漏洞。比如一个 PR 值极高的关于环保内容的大学页面，上面有一个链接连向一个儿童用品网站，这个链接出现的原因可能仅仅是因为这个大学页面维护人是个教授，他太太在那个卖儿童用品的公司工作。这种与主题无关、却有着极高 PR 值的链接，有可能使一些网站获得很好的排名，但其实相关性并不高。

Hilltop 算法就尝试矫正这种可能出现的疏漏。Hilltop 算法同样是计算链接关系，不过它更关注来自主题相关页面的链接权重。在 Hilltop 算法中把这种主题相关页面称为专家文件。显然，针对不同主题或搜索词有不同的专家文件。

根据 Hilltop 算法，用户搜索关键词后，Google 先按正常排名算法找到一系列相关页面并排名，然后计算这些页面有多少来自专家文件的、与主题相关的链接，来自专家文件的链接越多，页面的排名分值越高。按 Hilltop 算法的最初构想，一个页面至少要有两个来自专家文件的链接，才能返回一定的 Hilltop 值，不然返回的 Hilltop 值将为零。

根据专家文件链接计算的分值被称为 LocalRank。排名程序根据 LocalRank 值，对原本传统排名算法计算的排名做重新调整，给出最后排名。这就是前面讨论的搜索引擎排名阶段最后的过滤和调整步骤。

Hilltop 算法最初写论文和申请专利时对专家文件的选择有不同描述。在最初的研究中，Krishna Baharat 把专家文件定义为包含特定主题内容，并且有比较多导出链接到第三方网站的页面，这有点类似于 HITS 算法中的枢纽页面。专家文件链接指向的页面与专家文件本身应该没有关联，这种关联指的是来自同一个主域名下的子域名，来自相同或相似 IP 地址的页面等。最常见的专家文件经常来自于学校、政府及行业组织网站。

在最初的 Hilltop 算法中，专家文件是预先挑选的。搜索引擎可以根据最常见的搜索词，预先计算出一套专家文件，用户搜索时，排名算法从事先计算的专家文件集合中选出与搜索词相关的专家文件子集，再从这个子集中的链接计算 LocalRank 值。

不过在 2001 年所申请的专利中，Krishna Baharat 描述了另外一个挑选专家文件的方法，专家文件并不预先选择，用户搜索特定查询词后，搜索引擎按传统算法挑出一系列初始相关页面，这些页面就是专家文件。Hilltop 算法在这个页面集合中再次计算哪些网页有来自于集合中其他页面的链接，赋予比较高的 LocalRank 值。由于传统算法得到的页面集合已经具备了相关性，这些页面再提供链接给某一个特定页面，这些链接的权重自然应该很高。这种挑选专家文件的方法是实时进行的。

通常认为 Hilltop 算法对 2003 年年底的佛罗里达更新有重大影响，不过 Hilltop 算法是否真的已经被融入进 Google 排名算法中，没有人能够确定。Google 从来没有承认，也没有否认自己的排名算法中是否使用了某项专利。不过从排名结果观察及招揽 Krishna Baharat 至麾下等迹象看，Hilltop 算法的思想得到了 Google 的极大重视。

Hilltop 算法提示 SEO，建设外部链接时更应该关注主题相关的网站。最简单的方法是搜索某个关键词，目前排在前面的页面就是最好的链接来源，甚至可能一个来自竞争对手网站的链接效果是最好的。当然，获得这样的链接难度最大。

2.6 用户怎样浏览和点击搜索结果

用户搜索关键词后，搜索引擎通常返回 10 个结果。用户对这 10 个结果列表的浏览和点击有很大差别。本节介绍用户在搜索结果页面上的浏览方式，包括目光关注度及点击的一些研究。

2.6.1 英文搜索结果页面

页面浏览最主要的研究方法是视线跟踪（eye-tracking），使用特殊的设备跟踪用户目光在结果页面上的浏览及点击数据。enquiro.com 就是专门做这方面的实验及统计的公司。2005 年年初，enquiro.com 联合 eyetools.com 和 did-it.com 两家公司进行了一次很著名的视线跟踪实验，实验数据于 2005 年 6 月发表，提出在 SEO 业界很有名的金三角图像，也有人称其为“F 型”浏览图像，如图 2-30 所示。

图 2-30 中的颜色区块代表用户目光的停留位置及关注时间，图像中的 X 号代表点击。从图中我们可以看到，典型搜索用户打开搜索结果页面后，目光会首先放在最左上角，然后向正下方移动挨个浏览搜索结果，当看到感兴趣的页面时，横向向右阅读页面标题。排在最上面的结果得到的目光关注度最多，越往下越少，形成一个所谓的“金三角”。金三角中的搜索结果都有比较高的目光关注度。这个金三角结束于第一屏底部的排名结果，用户向下拉页面查看第二屏结果的概率大为降低。

这个浏览统计是针对 Google 搜索结果页面做的。后来 enquiro.com 针对雅虎及 MSN 搜索结果页面做的实验也得到大致相同的结果，如图 2-31 所示。

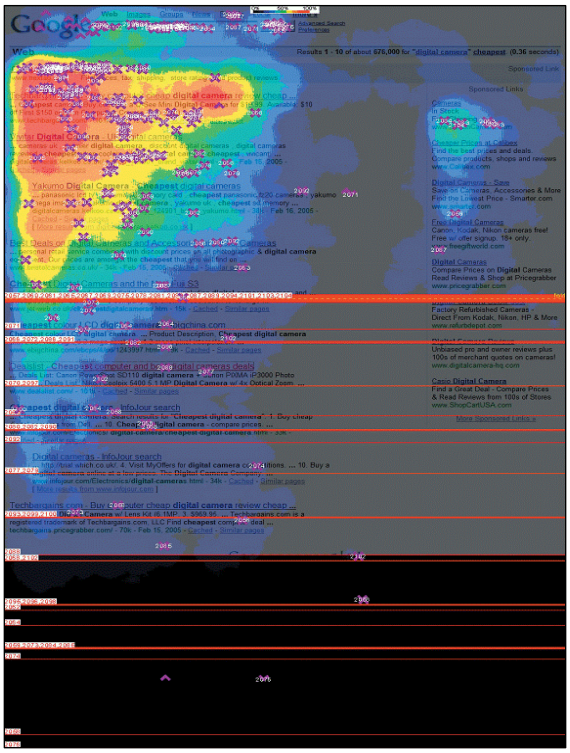


图 2-30　著名的用户视线分布金三角

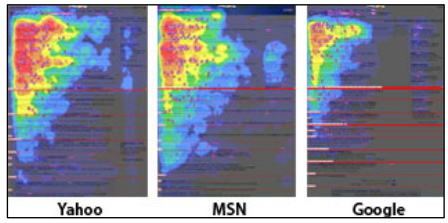


图 2-31　主流搜索引擎都存在视线分布金三角

2009 年 Google 官方博客也发布了一个类似的目光跟踪实验结果，确认了 enquiro.com 的金三角图像。Google 的实验结果如图 2-32 所示。

图 2-32　Google 官方发布的视线分布金三角

2006 年 10 月，康奈尔大学做了更进一步的实验和统计，记录 397 次实验对象（搜

索用户）对搜索结果的关注时间及点击分布，实验数据如图 2-33 所示。

图 2-33　康奈尔大学实验显示的搜索结果关注时间及点击分布

我们可以看到，排名前三位的页面得到的关注时间相差不大，尤其是前两位差距很小，但是点击次数却有很大差异。排名第一的结果占据了 56.36%的点击，排名第二的结果只有不到第一位四分之一的点击量，从第四位以后点击率更是急剧下降。唯一的特例是排名第十的点击结果，比第九位稍微多了一点。原因可能是用户浏览到最后一个结果时没有更多的结果可看，也没有其他选择，于是就点击了最后一个页面。

中间还有一个值得注意的结果，排名第七位的页面点击率非常低，只占 0.36%。这是因为前六位结果都处在第一屏，用户如果在第一屏没有找到满意结果，就会拉动右侧滑动条看第二屏内容。不过大多数用户不会刚好把屏幕下拉到第七位结果排在最上面，而是直接拉动到页面最下面，这样第七位结果反倒已经跑到第二屏之外，很多用户根本没看到第七位排名页面。

图 2-34 显示的是把关注时间及点击次数按曲线显示的结果。

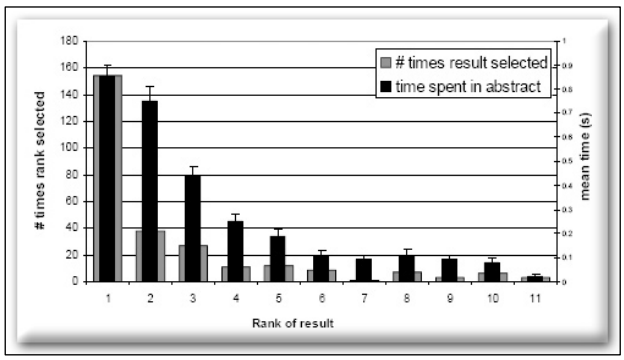


图 2-34　关注时间及点击次数按曲线显示

我们可以更清楚地看到，关注时间是按比较连续的曲线下降，而点击次数在第一位和第二位结果上有了巨大的差别，从第二位之后才形成比较平滑的曲线。

上面的实验数据来自于对用户搜索的观察记录，还不是来自搜索引擎的真正点击数据，而且样本数量有限。

2006 年 8 月，美国在线（AOL）因为疏忽公布了三个月的真实搜索记录，包括 2006 年 3 月 1 日到 5 月 31 日 1900 万次搜索，1080 多万不同的搜索词，还包括 65 万 8000 个用户 ID。这份资料公布后引起轩然大波。虽然用户 ID 都是匿名的，但是搜索词本身就可能泄露个人隐私，使得有心人士可以从这份资料中挖掘出不少能与具体个人相关联的数据。

也有 SEO 人士对这些搜索记录做了大量统计，得出搜索结果页面的真实点击数据。有人从 9 038 794 个搜索中统计到 4 926 623 次点击，这些点击在前 10 个结果中的分布如表 2-5 所示。

表 2-5　根据 AOL 泄露数据计算的搜索结果点击分布

页 面 排 名	点 击 次 数	占点击总数比例
1	2 075 765	42.1 %
2	586 100	11.90%
3	418 643	8.50%
4	298 532	6.10%
5	242 169	4.90%
6	199 541	4.10%
7	168 080	3.40%
8	148 489	3.00%
9	140 356	2.80%
10	147 551	3.00%

把表 2-5 中的点击比例换算成相对于第一位结果点击减少的百分比，又可以得到如表 2-6 所示的表格。

表 2-6　相对第一个结果的点击减少比例

页 面 排 名	点 击 次 数	比第一位点击减少倍数
1	2 075 765	
2	586 100	3.5
3	418 643	4.9
4	298 532	6.9
5	242 169	8.5

续表

页 面 排 名	点 击 次 数	比第一位点击减少倍数
6	199 541	10.4
7	168 080	12.3
8	148 489	14.0
9	140 356	14.8
10	147 551	14.1

从这个数据中可以看到，第一页点击分布与康奈尔大学的数据大体相当。排名第一的结果获得了 42.1%的点击率，排名第二的结果点击次数大幅下降，不到第一位的四分之一。排名第一页的 10 个结果，总共获得所有点击流量的 89.71%。第二页排名第 11～20 的结果，得到 4.37%的点击率。第三页只得到 2.42%。前 5 页占据了 99%以上的点击。

这是目前为止我们所能看到的唯一一份来自搜索引擎的真实点击数据，对 SEO 有很大的参考价值。

比如，同样是提高一位排名，从第十位提高到第九位，与从第二位提高到第一位获得的流量有天壤之别。很多公司和 SEO 人士把排名进入前十或前五当做目标，但实际上第十名或第五名与第一名流量上的差距非常大。这就给我们一个启示，有的时候我们可以找到网站有哪些关键词排名在第二位，想办法把它提高到第一位，能使流量翻好几倍。

这两个搜索结果点击数据，对 SEO 人员预估流量也有重要意义。

2.6.2 中文搜索结果页面

上面介绍的目光跟踪及点击数据，都是针对英文网站及美国用户的。那么中文搜索引擎情况如何呢？

2007 年 4 月，enquiro.com 做了 google.cn 及百度搜索结果页面实验。参加实验的是 50 个 18～25 岁的中国留学生，这些留学生来到美国不超过几个星期，正在就读语言培训班，所以其浏览习惯大体上还与主流中文用户相同，没有受英文用户浏览习惯太大的影响。这次实验的结果如图 2-35 所示。

图 2-35 是英文 google.com 与中文 google.cn 的对比。可以明显看出，相对于英文 Google 上比较规则的 F 型分布，中文用户在 google.cn 上的浏览更具随机性。虽然大体上还是呈现最上面的页面关注时间比较多，越往下越少，但是中文用户并不像英文用户那样垂直向下浏览结果，看到感兴趣的结果则向右方移动目光，阅读页面标题或说明。中文用户的目光更多地像是横向随机跳动，点击也是比较随机的，目光及点击分布都更广。

图 2-35　中文用户视线分布与英文用户对比

图 2-36 是百度和 google.cn 的搜索页面对比。如果说用户在 google.cn 上还大致符合越上面的页面关注越多的规律，在百度上则连在垂直方向也呈现更多随机特性，用户目光从上向下并没有显现出关注时间的急剧下降，百度用户不仅浏览页面上部结果，也在页面下部的结果上花了不少时间。在页面底部的相关搜索上，更是呈现出聚集目光和点击的情况。

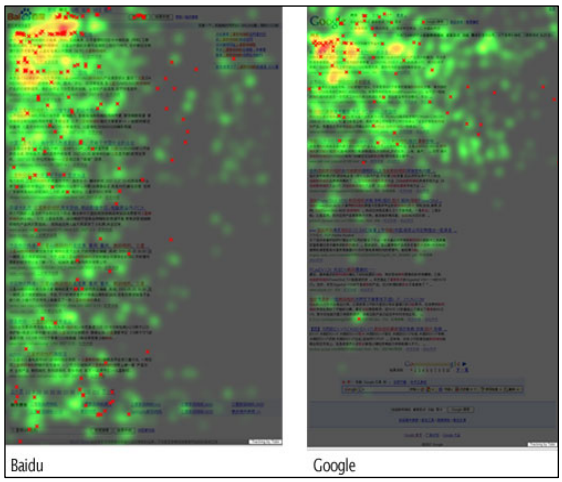


图 2-36　百度与 Google 中文的视线分布对比

按照英文用户搜索引擎结果浏览习惯分析，中文用户无论在 Google 还是百度上，似乎都花了更长时间才能找到自己想要的结果。英文用户在 Google 上平均 8～10s 就找到想要的结果，而中文用户在 google.cn 上则需要花 30s，在百度上要花 55s。这一方面说明中文搜索比英文搜索结果准确度低，另外也很可能是因为语言方面的差异。中文句子里的词都是连在一起的，用户必须花多一点时间真正阅读标题，才能了解列出的结果是否符合自己的要求。而英文单词之间有空格分隔，更利于浏览，用户很容易在一瞥之下就能看到自己搜索的关键词。

在百度上满天星似的浏览也可能与百度广告和自然结果都放在左侧、又没有背景颜色区分有关，一些用户会很自然地跳过广告，去查看排在后面的结果。

目前还没有见到中文搜索结果页面的点击数据统计。显然，前面介绍的点击数据不适用于中文搜索结果页面，尤其不适用于百度。可以想象，中文搜索结果点击率没有英文那样急剧下降的趋势，排在第五六位与排在第一位不会相差 10 倍之多。预估中文关键词流量时，不能照搬英文点击数据，而要更多地依靠自己网站的点击数据。

2.6.3 整合搜索及个人化搜索

上面的视线跟踪及点击实验，都是基于传统 10 个文字列表的搜索结果页面。近几年随着整合搜索和个人化搜索的流行，搜索结果页面中出现图片、视频、新闻等结果，整个页面排版方式的变化必然影响用户浏览方式。

2007 年 9 月，enquiro.com 又做了整合搜索的目光跟踪实验，结果如图 2-37 所示。

图 2-37　整合搜索结果对视线分布的影响

可以看到，当有图片出现在结果页面上部时，用户目光不再是从页面最左上角开始，而是首先把目光放在了图片上，接着向右移动看图片对应的结果是否符合自己的要求。然后用户视线再回到左上角重新向下浏览，看到合适的页面时再向右侧移动目光阅读页面标题和说明。

很明显，图片的出现完全改变了用户浏览方式，极大地吸引了用户目光。因此整合搜索结果不仅获得排名比普通页面要容易，竞争更小，而一旦出现排名也更能吸引用户眼球和点击。

enquiro.com 对整合结果的实验也发现，带有图片的列表常常起到一种分隔作用。

如图 2-38 所示，以图片为横向分隔线，分隔线之上的结果获得很大的目光关注。用户把图片当成了一个阻隔，视线不再向下浏览图片之下的内容。

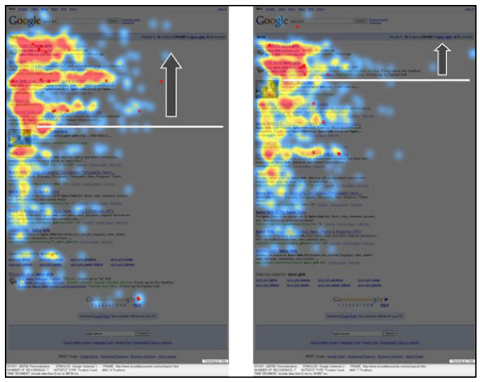


图 2-38　搜索结果中出现的图片起到分隔作用

这对传统关注于页面排名的 SEO 来说是个挑战，而且是自己无法克服的挑战。好在这个实验是 2007 年做的，当时整合搜索结果还是个新鲜事物，用户不太习惯，因此会吸引更多的不成比例的视线。当用户对带有图片、视频的结果习以为常后，很可能浏览方式会向传统金三角模式靠拢。

enquiro.com 同时做了个人化搜索页面的目光跟踪实验。

如图 2-39 所示，左侧是非个人化结果页面，右侧是个人化结果页面，白色框中出现的 3 个结果是个人化结果，也就是用户以前曾经访问过的网站。明显可以看到，用户对自己访问过的熟悉的网站投入的目光关注度和时间要比陌生的网址高得多。在个人化搜索一节我们会提到，虽然目前看到的个人化结果还很少，但这是一个确定的搜索行业趋势，对 SEO 的影响不仅仅在于排名，还在于高得多的点击率。

图 2-39　个人化结果对视线分布的影响

图 2-40 显示了非个人化搜索及个人化搜索的具体点击比例。

图 2-40 中左侧是非个人化搜索，右侧是个人化搜索，3 个数据分别为：

（1）所花时间比例；

（2）目光注视比例；

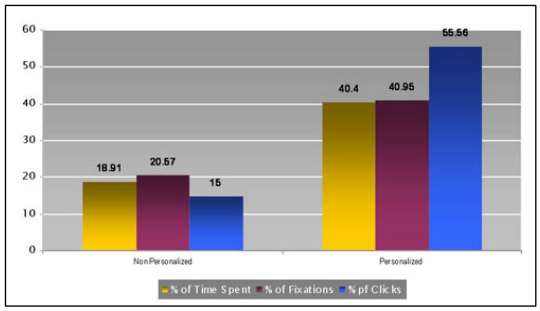


图 2-40　非个人化搜索及个人化搜索的点击比例

（3）点击比例。

我们可以看到，用户对熟悉的网站点击可能性要比不熟悉的网站高出四倍多。

2010 年 3 月，OneUpWeb 公司还做过实时搜索的视线跟踪实验，如图 2-41 所示，页面底部框中是实时搜索结果。

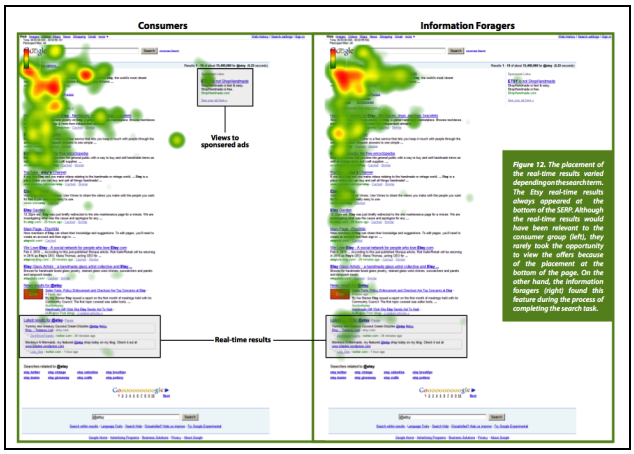


图 2-41　实时搜索对视线分布的影响

可以发现用户不太关注实时搜索，与整合搜索有比较大的差别。

2.7　高级搜索指令

用户除了可以通过搜索引擎搜索普通关键词外，还可以使用一些特殊的高级搜索指令。这些搜索指令对普通用户来说很少会用到，对 SEO 人员进行竞争对手研究和寻找外部链接资源却非常有用。本节简单介绍常用的高级搜索指令。

2.7.1　双引号

把搜索词放在双引号中，代表完全匹配搜索，也就是说，搜索结果返回的页面包含

双引号中出现的所有词，连顺序也必须完全匹配。百度和 Google 都支持这个指令。

比如：搜索“飞机微博图片”，如图 2-42 所示。

图 2-42　不带双引号的搜索结果

从图 2-42 中可以看到，返回的结果中大部分页面出现的关键词并不是完整的“飞机微博图片”，有的页面中“飞机”、“微博”、“图片”这三个词出现在不同地方，中间有间隔，顺序也不相同。把“飞机微博图片”放在双引号中再搜索“飞机微博图片”，如图 2-43 所示。

图 2-43　搜索时带双引号的结果

可以看到，返回结果只剩下 1 个，而且是完全按顺序出现“飞机微博图片”这个搜索字符串的页面。

使用双引号搜索可以更准确地找到特定关键词的竞争对手。

2.7.2　减号

减号（-）代表搜索不包含减号后面的词的页面。使用这个指令时，减号前面必须是空格，减号后面没有空格，紧跟着需要排除的词。Google 和百度都支持这个指令。

比如：搜索“新加”这个词时，返回结果如图 2-44 所示。

图 2-44　普通不使用减号的搜索

排在前面的很多是关于新加坡的页面。如果我们搜索“新加 –坡”，返回的则是包含“新加”这个词，却不包含“坡”这个词的结果，如图 2-45 所示。

图 2-45　搜索时使用减号的搜索结果

使用减号也可以更准确地找到需要的文件，尤其是某些词有多种意义时。比如，搜索“苹果 -电影”，返回结果页面就排除了《苹果》这部电影的结果，而不会影响苹果电脑和苹果作为水果的内容。

2.7.3　星号

星号（*）是常用的通配符，也可以用在搜索中。百度不支持*号搜索指令。

比如在 Google 中搜索“搜索*擎”，其中的*号代表任何文字。返回的结果不仅包含“搜索引擎”，还包含了“搜索收擎”、“搜索巨擎”等内容，如图 2-46 所示。

图 2-46　搜索时使用星号

2.7.4　inurl:

inurl: 指令用于搜索查询词出现在 URL 中的页面。百度和 Google 都支持 inurl:指令。inurl:指令支持中文和英文。

比如搜索“inurl:搜索引擎优化”，从图 2-47 中可以看到，返回的结果是网址 url 中包含“搜索引擎优化”的页面。由于关键词出现在 URL 中对排名有一定影响，因此，使用 inurl:搜索可以更准确地找到竞争对手。

图 2-47　inurl:指令

2.7.5　inanchor:

inanchor:指令返回的结果是导入链接锚文字中包含搜索词的页面。百度不支持 inanchor:指令。

比如，在 Google 搜索“inanchor:点击这里”，从图 2-48 中可以看到，返回的结果页面本身并不一定包含“点击这里”这四个字，而是指向这些页面的链接锚文字中出现了“点击这里”这四个字。

在后面的章节中我们会讨论，链接锚文字是现在关键词排名最重要的因素之一，有经验的 SEO 会尽量使外部链接锚文字中出现目标关键词。因此，使用 inanchor:指令可以找到某个关键词的竞争对手，而且这些竞争对手往往是做过 SEO 的。研究竞争对手页面有哪些外部链接，就可以找到很多链接资源。

图 2-48　inanchor:指令

2.7.6　intitle:

intitle: 指令返回的是页面 title 中包含关键词的页面。Google 和百度都支持 intitle:指令。

title 是目前页面优化的最重要因素。做 SEO 的人无论要做哪个词的排名，都会把关键词放进 title 中。使用 intitle 指令找到的文件才是更准确的竞争页面。如果关键词只出现在页面可见文字中，而没有出现在 title 中，大部分情况是并没有针对关键词进行优化，也不是有力的竞争对手。

比如搜索“搜索引擎优化”，本书第 1 版写作时（2010 年年初），笔者的博客在百度排在自然排名的第二位，如图 2-49 所示。

但是同一时间搜索“intitle:搜索引擎优化”，笔者的博客就不会被返回，因为当时博客标题中并没有“搜索引擎优化”这六个字，如图 2-50 所示。注意，“SEO 每天一贴”首页标题经常改变，有时候是因为做一些小实验，有时候是因为标题不再适用，如图 2-49 中所示的“正在写书”，书写后当然就修改了。所以读者看到本书时，博客标题肯定不一样了。

图 2-49　普通搜索结果

图 2-50　intitle:指令搜索结果

2.7.7　allintitle:

allintitle:搜索返回的是页面标题中包含多组关键词的文件。例如，

allintitle:SEO 搜索引擎优化

就相当于：

intitle:SEO intitle:搜索引擎优化

返回的是标题中既包含“SEO”，也包含“搜索引擎优化”的页面。

2.7.8　allinurl:

与 allintitle:类似。

allinurl:SEO 搜索引擎优化

就相当于：

inurl:SEO inurl:搜索引擎优化

2.7.9 filetype:

filetype 用于搜索特定的文件格式。Google 和百度都支持 filetype:指令。

比如：搜索“filetype:pdf SEO”返回的就是包含 SEO 这个关键词的所有 PDF 文件，如图 2-51 所示。

百度只支持下面几种文件格式：pdf，doc，xls，ppt，rtf，all。其中的“all”表示搜索百度所有支持的文件类型。Google 则支持所有能索引的文件格式，包括 HTML、PHP 等。

filetype:指令用来搜索特定的资源，比如 PDF 电子书、Word 文件等非常有用。

图 2-51 filetype:指令

2.7.10 site:

site:是 SEO 最熟悉的高级搜索指令，用来搜索某个域名下的所有文件。比如：搜索

site:seozac.com

返回的就是 seozac.com 这个域名下的所有页面，如图 2-52 所示。可见，这个指令是查询网站收录页面数的最直接的方法。

site:指令也可以用于子域名，比如：

site:blog.sina.com.cn

搜索的就是 blog.sina.com.cn 子域名下的所有收录页面。而

site:sina.com.cn

则包含 sina.com.cn 本身及 sina.com.cn 下面所有子域名（包括如 blog.sina.com.cn）下的页面。

不过 site:指令并不准确，尤其是 Google 中，返回的收录页面经常有大幅度波动，只能作为参考。

图 2-52　site:指令

2.7.11　link:

link:也是 SEO 常用的指令，用来搜索某个 url 的反向链接，既包括内部链接，也包括外部链接。比如搜索：

link: seozac.com

返回的就是 seozac.com 的反向链接，如图 2-53 所示。

不过可惜的是，Google 的 link:指令返回的链接只是 Google 索引库中的一部分，而且是近乎随机的一部分，所以用 link:指令查反向链接几乎没有用。百度则不支持 link:指令。

图 2-53　link:指令

2.7.12　linkdomain:

linkdomain:指令只适用于雅虎，返回的是某个域名的反向链接。雅虎的反向链接数据还比较准确，是 SEO 人员研究竞争对手外部链接情况的重要工具之一。比如：搜索

linkdomain:dunsh.org -site:dunsh.org

得到的就是点石网站的外部链接，因为-site:dunsh.org 已经排除了点石本身的页面，也就是内部链接，剩下的就都是外部链接了，如图 2-54 所示。

图 2-54　雅虎的 linkdomain:指令

2.7.13　related:

related:指令只适用于 Google，返回的结果是与某个网站有关联的页面。比如：搜索

related:dunsh.org

我们就可以得到 Google 所认为的与点石网站有关联的其他页面，如图 2-55 所示。

这种关联到底指的是什么，Google 并没有明确说明，一般认为指的是有共同外部链接的网站。

图 2-55　related:指令

2.7.14　综合使用高级搜索指令

对于前面介绍的几个高级搜索指令，单独使用可以找到不少资源，或者可以更精确地定位竞争对手。把这些指令混合起来使用则更强大。

比如下面这个指令：

inurl:gov 减肥

返回的就是 URL 中包含“gov”，页面中有“减肥”这个词的页面。很多 SEO 人员认为政府和学校网站有比较高的权重，找到相关的政府和学校网站，就找到了最好的链接资源。

下面这个指令返回的是来自.edu.cn，也就是学校域名上的包含“交换链接”这个词的页面：

inurl:.edu.cn 交换链接

SEO 人员从中可以找到愿意交换链接的学校网站。

或者使用一个更精确的搜索：

inurl:.edu.cn intitle:交换链接

返回的则是来自 edu.cn 域名，标题中包含“交换链接”这四个字的页面，返回的结果大部分应该是愿意交换链接的学校网站。

再如下面这个指令：

inurl:edu.cn/forum/*register

返回的结果是在.edu.cn 域名上，url 中包含“forum”以及“register”这两个单词的页面，也就是学校论坛的注册页面。找到这些论坛后，也就找到了能在高权重域名上留下签名的很多机会。

下面这个指令返回的是与减肥有关且 URL 中包含“links”这个单词的页面：

减肥 inurl:links

很多站长把交换链接页面命名为 links.html 等，所以这个指令返回的就是与减肥主题相关的交换链接页面。

下面这个指令返回的是 URL 中包含“gov.cn”和“links”的页面，也就是政府域名上的交换链接页面：

allinurl:gov.cn+links

最后一个例子，通过雅虎搜索这个指令：

linkdomain:dunsh.org -linkdomain:seozac.com

返回的是链接到点石网站，却没有链接到我的博客的网站。使用这个指令可以找到很多链向你的竞争对手或其他同行业网站、却没链向你的网站的页面，这些网站是最好的链接资源。

高级搜索指令组合使用变化多端，功能强大。一个合格的 SEO 必须熟练掌握这几个常用指令的意义及组合方法，才能更有效地找到更多的竞争对手和链接资源。

第 3 章
竞争研究

初做网站的人很容易犯的最大错误之一是：脑袋一拍就贸然进入某个领域，跳过竞争研究，没规划好目标关键词就开始做网站。这样做常常导致两个结果，一是，自己想做的关键词排名怎么也上不去，二是，自己认为不错的关键词排名到了第一也没什么流量。

进行竞争研究，确定适当的关键词是 SEO 的第一步，而且是必不可少的一步。竞争研究包括关键词研究、竞争对手研究及现有网站评估诊断，其中关键词研究是最重要的。

3.1 为什么要研究关键词

研究关键词的意义在于以下几方面。

3.1.1 确保目标关键词有人搜索

网站核心关键词不能想当然地选用，必须经过关键词研究才能确保这个关键词确实有网民在搜索，没人搜索的词没有任何价值。

对 SEO 没概念的人确定目标关键词时常常会首先想到公司名称或自己的产品名称。但是当企业或网站没有品牌知名度时，没有用户会搜索公司名或网站名。产品名称如果不包含产品的通用名称，也往往没人搜索。

很多时候，即使使用行业最通用的名字，也不一定有足够的真实搜索次数。最典型的就是“SEO”这个词本身。百度指数显示“SEO”每天被搜索两三千次以上，首先这与很多其他真正的用户在搜索的词相比，已经属于比较低的。其次，搜索“SEO”的人很多是站长在研究记录“SEO”这个词排名的新动向，而不是对 SEO 服务感兴趣。

要确定适当的关键词，首先要做的是，确认用户搜索次数达到一定数量级。如果在这方面做出错误的方向选择，对网站的影响将会是灾难性的。

3.1.2 降低优化难度

要找到有搜索量的关键词，并不意味着就要把目标定在最热门、搜索次数最多的词上。虽然搜索“新闻”、“律师”、“租房”、“机票”、“减肥”、“旅游”、“化妆品”等这些词的用户很多，但是对中小企业和个人站长来说，要把这些词做到前几位，难度非常大。可以说，没有强大的资源、人力支持，想都不用想。做关键词研究就是要找到被搜索次

数比较多，同时难度不太大的关键词，网站优化才有可能在一定的预算、周期下取得较好的效果。

3.1.3 寻找有效流量

排名和流量都不是目的，有效流量带来的转化才是目的。就算公司有足够的实力将一些非常热门的关键词排到前面，也不一定是投入/产出比最好的选择。

假设网站提供律师服务，将核心关键词定为“律师”，一般来说并不是好的选择，因为搜索“律师”的用户动机和目的是什么很难判定。用户有可能在寻找律师服务，但也可能是在寻找律师资格考试内容，还可能是在找大学专业报考指导，这样的用户来到提供律师服务的网站就没有什么机会转化为付费客户。

如果把核心关键词定为“北京律师”，针对性就要强得多，用户已经透露出一定的购买意向。再进一步，如果目标关键词定为“北京刑事律师”，则购买意向或者说商业价值就更高，几乎可以肯定这个用户是在寻找特定区域、特定案件的律师服务，这样的搜索用户来到你的网站，转化为客户的可能性将大大提高。

要记住，流量本身并不一定是资产，很可能是浪费带宽、客服的无谓付出，只有能转化的有效流量才是资产。

3.1.4 搜索多样性

搜索词并不局限于我们容易想到的热门关键词。用户使用的搜索词五花八门，很多是站长自己所想象不到的。

据著名市场调研公司 Hitwise 于 2009 年 2 月的研究显示，用户实际使用的搜索词有越来越长的趋势。如表 3-1 所示是 2008 年 1 月到 2009 年 1 月用户搜索关键词长度变化趋势。

表 3-1　用户搜索词变长

关键词点击比例				
主　　题	2008 年 1 月	2008 年 12 月	2009 年 1 月	年比变化
1 个单词	20.96%	20.70%	20.29%	–3%
2 个单词	24.91%	24.13%	23.65%	–5%
3 个单词	22.03%	21.94%	21.92%	0
4 个单词	14.54%	14.67%	14.89%	2%
5 个单词	8.20%	8.37%	8.68%	6%
6 个单词	4.32%	4.47%	4.65%	8%
7 个单词	2.23%	2.40%	2.49%	12%
8 个单词及以上	2.81%	3.31%	3.43%	22%

可以看到，搜索 3 个单词以下关键词的比例不断下滑，搜索 4 个词以上的则上升，词越长，上升得越多。Hitwise 的研究虽然针对的是英文搜索，但是没有理由认为中文搜索不是同样的趋势。

随着搜索经验越来越丰富，网民已经知道搜索很短的、一般性的词，往往找不到自己想要的内容，而搜索更为具体的、比较长的词效果更好。做过网站的人都会从流量分析中发现，很多用户现在不仅搜索关键词，甚至搜索完整的句子。

无论是从用户意图和商业价值看，还是从搜索词长度来看，更为具体的、比较长的搜索词都有非常重要的意义。SEO 人员必须知道，用户除了搜索行业通称，还在搜索哪些更具体的词、搜索次数是多少，这样才能确定网站核心关键词。

3.1.5 发现新机会

每个人的思维都会有局限。研发和销售某些特定产品的人，思路很容易被局限在自己和同事最常用的词汇上。而用户需求千变万化，上网经验也不同，他们会搜索各种各样我们想不到的词。

SEO 人员查询搜索引擎提供的关键词扩展工具，或者分析网站流量是一件非常有意思，而且常常有意外发现的工作。我个人的经验是，利用关键词工具的推荐，挖掘相关关键词，很容易几个小时就过去了，能看到太多自己完全不会去搜索的词汇，但却实实在在有用户在搜索。在这个过程中，经常发现有共通性或明显趋势的词，把这些关键词融入到网站上，甚至增加新栏目，是发现新机会、拓展内容来源的最好方式之一。

3.2 关键词的选择

选择恰当的关键词是 SEO 最具技巧性的环节之一。只有选择正确的关键词，才能使网站 SEO 走在正确的大方向上。确定什么样的关键词决定了网站内容规划、链接结构、外部链接建设等重要的后续步骤。

在介绍选择关键词步骤前，我们先讨论选择关键词的原则。

3.2.1 内容相关

目标关键词必须与网站内容有相关性。SEO 早期曾经流行在页面上设置甚至堆积搜索次数多但与本网站没有实际相关性的关键词，也起到过很好的作用，能带来不少流量。现在这样的做法早已过时。网站需要的不仅仅是流量，更是有效流量，可以带来订单的流量。靠欺骗性的关键词带来访客却不能完成转化，对网站毫无意义。这样的排名和流量不是资产，而是负担，除了消耗带宽，没有其他作用。

如果你的网站想卖“上海律师服务”或者“特效减肥茶”，就不要想着靠“世博会”

或“刘德华”这种关键词带来流量。抛开难度和可能性不谈，就算搜索这种不相关关键词的访客来到网站，也不会买你的产品或服务。

当然，这不一定适用于某些网站，比如，新闻门户或纯粹依靠广告赢利的信息类网站。很多门户类网站包罗万象，内容相关性判断也比较模糊。对这些网站来说，只要有流量，就有一定的价值，网站并不依靠本身的转化赢利。

3.2.2 搜索次数多，竞争小

很显然，最好的关键词是搜索次数最多、竞争程度最小的那些词，这样既保证 SEO 代价最低，又保证流量最大。可惜现实不是这么理想的。大部分搜索次数多的关键词，也是竞争大的关键词。不过，通过大量细致的关键词挖掘、扩展，列出搜索次数及竞争程度数据，还是可以找到搜索次数相对多、竞争相对小的关键词。

研究搜索次数比较直接、简单，Google 关键词工具及百度指数都提供搜索次数数据，详细内容请参考 SEO 工具相应章节。

竞争程度的确定比较复杂，需要参考的数据较多，而且带有比较大的不确定性。这部分内容请参考后面“关键词竞争程度判断”一节。

根据搜索次数和竞争程度可以大致判断出关键词效能。在相同投入的情况下，关键词效能高的词获得好排名的可能性较大，可以带来更多的流量。

3.2.3 主关键词不可太宽泛

这实际上是上面两点的自然推论。关键词宽泛，竞争太大，所花代价太高，搜索词意图不明确，转化率也将降低。做房地产的公司，想当然地把“房地产”作为目标关键词，做旅游的公司就把“旅游”作为目标关键词，这都犯了主关键词过于宽泛的毛病。

一般行业通称都是过于宽泛的词，如“新闻”、“旅行”之类。把目标定在这种宽泛的词上，要么做不上去，要么费了九牛二虎之力做上去却发现转化率很低，得不偿失。

3.2.4 主关键词也不可太特殊

选择主关键词也不能走向另一个极端。太特殊、太长的词，搜索次数将大大降低，甚至没有人搜索，不能作为网站的主关键词。

如果说“律师”这个词太宽泛，那么选择“北京律师”比较适当。根据不同公司业务范围，可能“北京刑事律师”更合适。但是如果选择“北京新街口律师”就不靠谱了。这种已经属于长尾的关键词可以考虑以内页优化，放在网站首页肯定不合适。

太特殊的关键词还包括公司名称、品牌名称、产品名称等。

所以，网站主关键词（或者称为网站核心关键词）既不能太长、太宽泛，也不能太短、太特殊，需要找到一个平衡点。

3.2.5 商业价值

不同的关键词有不同的商业价值，就算长度相同，也会导致不同的转化率。

比如，搜索“液晶电视原理”的用户购买意图就比较低，商业价值也低，他们很可能是在做研究，学习液晶电视知识而已。而搜索“液晶电视图片”的用户商业价值有所提高，很可能是在寻找、购买液晶电视的过程中想看看产品实物是什么样。搜索“液晶电视价格”，购买意图大大提高，已经进入产品比较选择阶段。而搜索“液晶电视促销”或“液晶电视购买”，其商业价值进一步提高，一个大减价信息就可能促成用户做出最后的购买决定。

在做关键词研究时，SEO 人员可以通过各种方式查询到大量的搜索词，通过常识就能判断出不同词的购买可能性。购买意图强烈、商业价值较高的关键词应该是优化时最先考虑的，无论内容规划，还是内部链接安排，都要予以侧重。

3.3 关键词竞争程度判断

关键词选择最核心的要求是搜索次数多、竞争程度小。搜索次数可以通过搜索引擎本身提供的关键词工具查看，简单明了，数字比较确定。

而竞争程度判断起来就要复杂得多。下面列出几个可以用于判断关键词竞争程度的因素。每个因素单独看都不能完整、准确地说明关键词的竞争情况，而必须整体考虑。更为困难的是，有的因素在数值上并不确定，比如竞争对手网站优化水平，无法给出一个确定数值。这几个竞争程度表现因素哪个占的比例更大，也没有一定的结论。所以，基于经验的主观判断就变得非常重要了。

3.3.1 搜索结果数

搜索结果页面右上角都会显示这个关键词返回的相关页面总数。这个结果数是搜索引擎经过计算认为与搜索词相关的所有页面，也就是参与这个关键词竞争的所有页面。

显然，结果数越多，竞争程度越大。通常，结果数值在 10 万以下，竞争很小，稍微认真地做一个网站，就可以获得很好的排名。权重高的域名上经过适当优化的内页也可以迅速获得排名。

结果数达到几十万，说明关键词有一定难度，需要一个质量和权重都不错的网站才能竞争。

结果数达到一两百万以上，说明关键词已经进入比较热门的门槛。新网站排名到前几位的可能性大大降低，需要坚持扩展内容，建立外部链接，达到一定域名权重才能成功。

结果数达到千万级别以上，通常是行业通用名称，竞争非常激烈，只有大站、权重高的网站才能获得好的排名。

上面只是大致而论，实际情况千差万别。有的关键词虽然结果数很大，但没有任何商业价值，竞争程度并不高，比如，“我们”、“方法”、“公园”等这些常见却很一般化的词，出现在网页上的概率高，搜索结果数自然很高，但搜索的次数不多，商业价值很小。

而某些关键词看似结果数并不多，但是因为商业价值高，竞争程度非常高。比如，某些疑难杂症治疗方法、药品等。

3.3.2 intitle 结果数

使用 intitle：指令搜索得到的结果页面数如图 3-1 所示。

单纯搜索关键词返回的结果中（包括页面上出现关键词，但页面标题中没有出现的页面），这些页面虽然也有一点相关性，但很可能只是偶然在页面上提到关键词而已，并没有针对关键词优化，这些页面针对这个特定关键词的竞争实力很低，在做关键词研究时可以排除在外。

图 3-1　百度 intitle:指令搜索结果数

标题中出现关键词的页面才是真正的竞争对手。

3.3.3 竞价结果数

搜索结果页面右侧以及最上和最下面有多少个广告结果，也是衡量竞争程度的指标之一。

一般来说，广告商内部有专业人员做关键词研究和广告投放，他们必然已经做了详细的竞争程度分析及赢利分析，只有能产生效果和赢利的关键词，他们才会去投放广告。如果说搜索结果数还只是网上内容数量带来的竞争，竞价数则是拿着真金白银与你竞争的真实存在的竞争对手数目。

搜索结果页面右侧广告最多显示 8 个，比较有商业价值的关键词，通常都会显示满 8 个广告结果。如果某个关键词搜索页面右侧只有两三个广告，说明关注这个词的网站还比较少，竞争较低。

要注意的是，竞价数需要在白天工作时间查看。广告商投放竞价广告时，经常会设置为晚上停止广告。对于经常晚上工作的 SEO 人员来说，如果半夜查看搜索结果页面，没看到几个广告商，就认为没有多少人参与竞价，很可能导致误判。

3.3.4 竞价价格

几大搜索引擎都提供工具，让广告商投放前就能看到某个关键词的大致价格，能排到第几位，以及能带来多少点击流量。如 Google 的关键词流量估算工具：

https://adwords.google.cn/select/TrafficEstimatorSandbox

如图 3-2 所示是 Google 流量估算工具显示的与“减肥”相关的关键词竞价价格及预估点击流量。

流量估算工具

« 修改设置 | 以 csv 格式下载

所有估算值仅作为参考指南提供，它们是根据整个系统内的平均值计算的。您的实际费用和广告排名可能有所不同。要查看根据您关键字的效果历史记录计算的估算值，请使用相应广告组中的流量估算工具。 了解详情

平均每次点击费用: ¥2.49 (最高每次点击费用为 ¥3.82)
估算的每日点击次数: 57 - 72 (每日预算为 ¥200.00)

估算值取决于您的出价以及地理定位选择。因为流量估算工具不会考虑您的每日预算，您的广告获得的点击次数可能低于估算值。

最高每次点击费用: 每日预算 获得新的估算值

关键字 ▼	搜索量	估算平均每次点击费用	估算的广告排名	估算的每天点击量	估算的每天费用
[减肥]		¥1.75 - ¥2.75	1 - 3	36 - 46	¥70 - ¥200
[减肥方法]		¥1.15 - ¥1.81	1 - 3	6 - 8	¥10 - ¥20
[健康减肥]		¥1.80 - ¥2.62	1 - 3	2	¥10
[快速减肥]		¥2.10 - ¥2.83	1 - 3	7 - 9	¥20 - ¥30
[瘦身]		¥2.20 - ¥3.00	1 - 3	6 - 7	¥20 - ¥30
搜索网络总计		**¥1.77 - ¥2.67**	**1 - 3**	**57 - 72**	**¥200**

« 修改设置 下载为 .csv 文件

图 3-2 Google 流量估算工具显示相关的关键词竞价价格及预估点击流量

显然，竞价价格越高，竞争程度也越高。当然也不能排除两三个广告商为了争抢广告位第一名而掀起价格战，把本来竞争程度不太高的关键词推到很高价格。实际上，如果广告商只想出现在第四、第五位的话，竞价价格大幅下降，参与竞价的广告商人数也没有那么多。

另外，某些利润率高的行业，搜索竞价经常超出自然排名的真正竞争程度，比如律师服务、特效药品，这些行业产品及服务的特性决定了一个订单的利润可能是成千上万，企业可以把竞价提到相当高的程度，甚至一个点击几十元钱也不罕见。而销售书籍、服装、化妆品等，利润不是很高，竞价价格也不可能太高。广告价格的巨大差异，并不能说明自然搜索竞争程度真的有这么大差别。

3.3.5 竞争对手情况

自然结果排在前面的主要竞争对手情况，包括外部链接数量质量、网站结构、页面关键词优化等。这部分很难量化，而且本身包含了众多因素。我们在后面“竞争对手网站研究”一节再详细讨论。

3.3.6 内页排名数量

搜索结果页面前 10 或前 20 位中，有多少是首页？有多少是网站内页？这在一定程度上说明了竞争水平。一般来说，排在前面的内页数越多，说明竞争越小。

通常，网站首页是权重最高的页面，排名能力也最强。如果一个关键词排在前 20 位的多数是网站内页，说明使用首页特意优化这个关键词的网站不多。如果自己的网站首页针对这个关键词，获得好排名的机会就比较大。如果有权重比较高的域名，分类页面甚至产品页面也都有机会。

要注意的是，这里所说的内页指的是一般网站的内页。如果排在前面的有很多大型知名门户的频道首页，这种内页实际上应视同网站首页。这种权重高的网站，频道首页权重也比一般网站首页高得多。

综合上面几种指标，SEO 人员可以给关键词打一个竞争程度指数，如表 3-2 所示（以 Google 搜索结果为例）。

表 3-2　关键词竞争程度指数示例

关键词	搜索结果数	intitle 数	竞价数	平均点击价格	竞争对手实力	前 10 内页数	竞争指数
减肥	66 400 000	9 030 000	8+3+2	￥1.64	9	4	10
运动减肥	17 800 000	829 000	6+1+1	￥0.99	5	8	7
快速减肥	16 400 000	2 050 000	8+3+2	￥2.04	7	4	9
饮食减肥	13 900 000	6 200 000	5+0+0	￥0.94	6	7	8
节食减肥	1 870 000	131 000	1+0+0	0	5	9	4
苹果减肥	3 370 000	5 790 000	1+0+0	￥0.76	6	8	7
腹部减肥	2 060 000	278 000	3+0+0	￥7.19	5	8	7
脸部减肥	6 520 000	192 000	1+0+0	0	3	10	3
快走减肥	5 110 000	14 500	0	￥0.34	4	9	3
跑步减肥	1 460 000	273 000	0	￥1.07	4	9	6

注：表中竞价数 8+3+2 是指搜索结果页面右侧广告 8 个，上部（自然搜索结果前）3 个，下部 2 个。竞争对手实力的判断请参考 3.11 节。

3.4 核心关键词

选择关键词的第一步是确定网站核心关键词。

核心关键词通常就是网站首页的目标关键词。一般来说，整个网站会有很多目标关键词，这些关键词不可能都集中在首页上进行优化，而是合理地分布在整个网站，形成金字塔形结构。难度最高、搜索次数最多的两三个是核心关键词，放在首页；难度次一级、数量更多的关键词，放在栏目首或分类首页；难度更低的关键词，数量更为庞大，放在具体产品或文章页面。

整个网站的关键词按照搜索次数、竞争程度、优化难度逐级分布。关键在于确定核心关键词，首页核心关键词一旦确定，其下的栏目及产品页面关键词也就相应确定了。

3.4.1 头脑风暴

确定核心关键词第一步，是列出与自己网站产品相关的尽量多的、同时比较热门的搜索词，自己先可以通过头脑风暴列出待选词。

建议问自己如下几个问题：

- 你的网站能为用户解决什么问题？
- 用户遇到这些问题时，会搜索什么样的关键词？
- 如果你自己是用户，在寻找这些问题的答案时会怎么搜索？
- 用户在寻找你的产品时会搜索什么关键词？

只要具备一定的常识并且了解自己的产品，就一定会列出至少一二十个备选核心关键词。

3.4.2 同事、朋友

一个人的思路有限，可以找几个同事一起“头脑风暴”。不要给自己设限，想到什么就全都记下来。

有的时候公司内部人员对自己的产品太过熟悉，反倒限制了思路，不容易从普通用户角度出发。可以问一下公司之外的亲戚朋友，要寻找你公司的产品或服务时会搜索什么关键词。

3.4.3 竞争对手

另一个备选关键词来源是竞争对手。查看竞争对手网站首页源文件，关键词标签列

出了什么关键词？标题标签中又出现了什么关键词？有实力的竞争对手应该已经做了功课，他们的网站所针对的关键词很可能就是不错的选择，如图 3-3 所示。

```
<html>
<head><title>减肥方法,快速减肥,健康减肥,[illegible]欢迎您!</title>
<meta http-equiv=Content-Type content="text/html; charset=gb2312">
<meta content=3mdMplJ6LDsAgnz0c+ajhpak+kpFXZfQyueBEUj4yCA= name=verify-v1>
<meta content=减肥方法,快速减肥,减肥瘦身,如何减肥,健康减肥,运动减肥,减肥药,减肥茶
name=keywords>
<meta content=[illegible]是专业的减肥方法网站,为朋友们提供减肥方法,快速减肥,减肥食谱,减
肥操,明星减肥,健康减肥,减肥药,减肥产品等资讯.告诉大家如何健康减肥方法,希望在我们的帮助
下,渴望快速减肥的朋友拥有健康苗条的好身材! name=description>
```

图 3-3　竞争对手网站元标签

现在很多网站没有写关键词标签。标题标签长度有限，再加上吸引点击的考虑，不一定完整列出目标关键词。但首页正文内容中一定会出现网站核心关键词。

Google 关键词工具：

https://adwords.google.cn/select/KeywordToolExternal

有一个功能很多人都忽略了，除了列出关键词搜索次数外，Google 工具还可以根据某个页面正文内容提炼出最相关的关键词。如图 3-4 所示，在“您希望如何生成关键字提示”下选择“网站内容”，在右侧输入竞争对手网址。

图 3-4　在 Google 关键词工具选择按“网站内容”生成关键词

Google 工具根据页面内容生成相关关键词，如图 3-5 所示。

无论网站怎么写标签，正文内容是没办法隐藏的，从正文中提炼的关键词通常是很准确的。

3.4.4　查询搜索次数

经过自己及朋友、同事的头脑风暴和检查竞争对手网站之后，再到 Google 关键词工具或百度指数工具，查询这些关键词的搜索次数。同时，这两个工具也会提供很多相关关键词，也许有前面没有想到也没有查到的。

Google 关键词工具或百度指数使用详情见 SEO 工具章节。

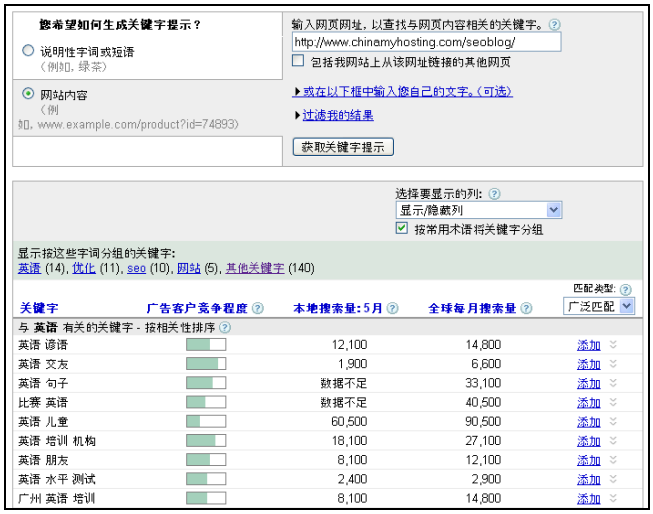


图 3-5　Google 关键词工具根据页面内容生成的相关关键词

选出其中搜索次数比较多的几十个关键词，记录下 Google 搜索次数、百度指数，再列出这些关键词对应的竞争指数，如表 3-3 所示。

表 3-3　关键词对应的竞争指数及潜在效能示例

关 键 词	Google 月搜索次数	百 度 指 数	竞 争 指 数	潜 在 效 能
减肥	673 000	7645	10	7
运动减肥	201 000	386	7	6
快速减肥	301 000	803	9	5
饮食减肥	135 000	118	8	4
节食减肥	165 000	155	4	8
苹果减肥	8100	176	7	5
腹部减肥	5400	202	7	5
脸部减肥	60 500	163	3	7
快走减肥	1000	0	3	4
跑步减肥	1300	236	6	2

有时候需要快速查看几个关键词搜索次数对比，可以在 Google 趋势和百度指数输入多个关键词，之间用逗号隔开，这两个工具将直观显示搜索量差别，如图 3-6 和图 3-7 所示。

3.4.5　确定核心关键词

面对这几十个甚至可能上百个关键词，要选出两三个作为网站核心关键词，通常有几种情况和策略。

图 3-6　Google 趋势直观显示搜索量差别

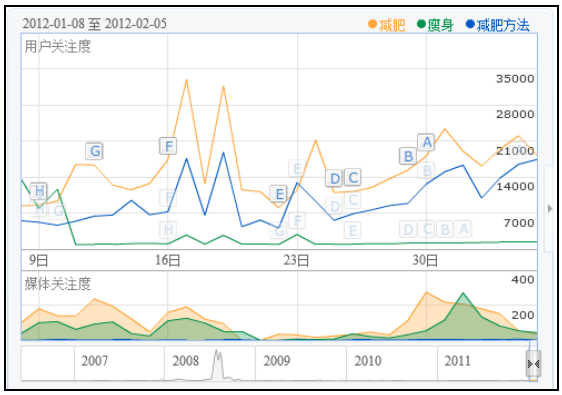


图 3-7　百度指数直观显示搜索量差别

对中小企业网站、个人网站及有志于电子商务的新站来说，核心关键词最好是效能最高的几个关键词，也就是搜索次数相对比较多、竞争指数相对比较小的几个，如表 3-3 中的“节食减肥”、“脸部减肥”。这样既保证了足够的搜索量及优化成功后带来的流量，又兼顾可行性。新站、小站把目标放在搜索次数最多的热门关键词上是不太现实的。

对于有资源、有实力并且有决心的公司来说，也可以把目标定在搜索次数最多的几个关键词上，只要这几个词不是太宽泛。像前面所说，除非你想打败新浪、搜狐，不然把“新闻”定为核心关键词是毫无意义的。就算把核心关键词确定在搜索次数最多的几个词上，也要做关键词研究，因为次一级关键词在网站上的分布还是需要研究和安排的。

有的公司产品早就存在，没有太大的灵活性，也不得不把核心关键词放在产品通用称呼上。这时候可能需要在产品名称前加限定词，才有可能获得比较好的排名。

注：本节中列出的关键词搜索次数、竞争指数、潜在效能等只是示例，请读者不要解读为 SEO 建议。主要有以下三点原因：

（1）作为示例，并没有列出可以判断全局的足够多的关键词。

（2）各公司或站长有自己的产品情况，不能仅看关键词，还要与自己的产品相关。例如，如果公司销售运动器械，显然不适合把“节食减肥”作为主要目标。

（3）笔者本人对减肥行业没有任何了解，标识的竞争指数可能与实际情况有很大出入。目前排名靠前的网站可能是减肥行业很强的域名，笔者不了解，只作为普通网站判断。

无论采取哪种策略，一般来说核心关键词在三个左右比较合适，不要超过 4 个。

上面过程列出的其他关键词数据，后面还会用到，并不会浪费。

3.5 关键词扩展

确定了核心关键词后，接下来就是进行关键词扩展。对一个稍有规模的网站来说，研究几十个关键词并不足够，还需要找出比核心关键词搜索次数少一些的更多关键词，安排到分类或频道首页。挖掘扩展出几百甚至几千个关键词都很常见。

扩展关键词可以通过下面几种方式。

3.5.1 关键词工具

最常见的工具还是 Google 关键词工具及百度指数。尤其 Google 关键词工具在这方面更方便，查询任何一个关键词，Google 工具都会列出至少几十个相关关键词。再取其中的任何一个重新查询，又可以带出另外几十个关键词。通过这种联想式的不断挖掘，扩展几百几千个关键词轻而易举，如图 3-8 所示。

SEO 人员可以把 Google 给出的关键词下载为 Excel 文件，在 Excel 中合并得到所有相关词，删除重复及搜索量很低的关键词，按搜索次数排序，就得到有一定搜索量的大量关键词。

明星 减肥	74,000	40,500	添加
减肥 食品	74,000	33,100	添加
运动 减肥 计划	数据不足	590	添加
减肥 网站	1,900	3,600	添加
			添加所有 150 个关键字？
		下载所有关键字：	文本，.csv(Excel 文件)，.csv
可供参考的其他关键字 - 按相关性排序			
jianfei	5,400	9,900	添加
节食	12,100	12,100	添加
快速 瘦 身 法	260	590	添加
减肥	12,100	165,000	添加
瘦 身 方法	165,000	90,500	添加
快速 瘦 身	33,100	33,100	添加
瘦 身 法	5,400	9,900	添加
苹果 瘦 身	数据不足	880	添加
瘦 身 食谱	5,400	6,600	添加
瘦 身 运动	74,000	12,100	添加
瘦 身 汤	8,100	12,100	添加
瘦 身 水果	12,100	9,900	添加
瘦 身 健康	9,900	12,100	添加
瘦 身 餐	数据不足	2,400	添加
瘦 身	2,240,000	1,500,000	添加
苦瓜 瘦 身	数据不足	880	添加
肥胖	74,000	74,000	添加
塑身	823,000	550,000	添加
七 日 瘦 身 汤	3,600	4,400	添加

图 3-8　Google 关键词工具生成大量相关关键词

百度指数查询时也会显示相关关键词，如图 3-9 所示。

图 3-9　百度指数显示的相关关键词

3.5.2　搜索建议

在百度或 Google 搜索框中输入核心关键词，搜索框会自动显示与此相关的建议关键词，如图 3-10、图 3-11 所示。

图 3-10　百度搜索建议

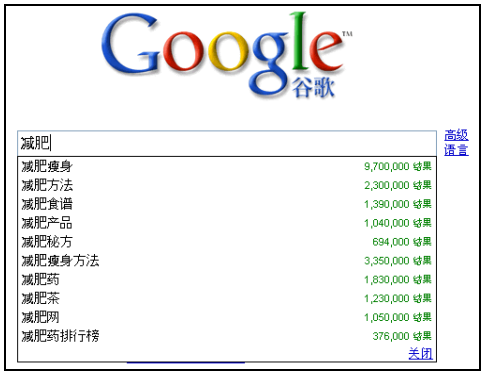


图 3-11　Google 搜索建议

3.5.3　相关搜索

搜索结果页面最下面可以看到搜索引擎给出的相关搜索，如图 3-12、图 3-13 所示。

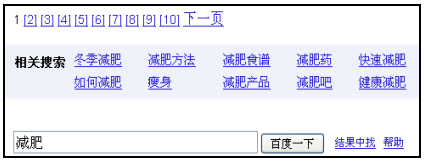


图 3-12　百度相关搜索

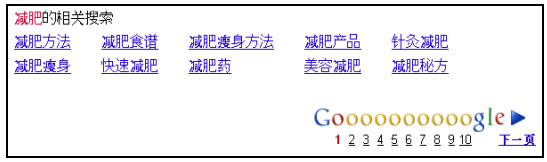


图 3-13　Google 相关搜索

一般来说，搜索建议和相关搜索中的扩展词在 Google 关键词工具中都会出现，但搜索建议和相关搜索使用最简单，是快速开拓思路的好方法。

3.5.4　其他关键词扩展工具

如后面 SEO 工具部分所介绍的追词网，也可以列出大量相关关键词，并显示百度指数，如图 3-14 所示。

图 3-14　追词网生成的相关关键词

3.5.5　各种形式的变体

1．同义词

假设核心关键词是酒店，与酒店基本同义的还有饭店、旅馆、住宿、旅店、宾馆等。再如网站推广、网络推广、网络营销意义也很相近。

2．相关词

虽然不同义，作用却非常类似的词。如网站建设、网页设计、网络营销与 SEO 非常相关，目标客户群也大致相同。

3．简写

Google PR 与 Google PageRank，北京大学与北大。

4．错字

还有一类变体是错字。如艾滋病与爱滋病，SEO 每天一贴与 SEO 每天一帖，点石与电石。有不少用户使用拼音输入法经常会输入错字、同音字，所以有一些搜索量。但是优化错别字就不可避免地要在页面中出现这些错别字，可能会给网站用户带来负面观感，使用时需要非常小心。

3.5.6 补充说明文字

核心关键词可以加上各种形式的补充说明。

1．地名

有的核心关键词配合地名很明显，如：

- 旅游——云南旅游，海南旅游，北京旅游。
- 酒店——上海酒店，南京酒店。

有的关键词看似与地理位置无关，却有不少用户会加上地名搜索，如：

- 北京办公家具，上海办公家具，广州办公家具。
- 上海鲜花，深圳鲜花。

哪种核心关键词配合地名有搜索量，需要查看搜索引擎关键词工具。有的关键词就完全没有人加上地名搜索，如减肥。

2．品牌

核心关键词加上品牌名称也很常见，如：

- 电视机——康佳电视机，索尼电视机。
- 手机——诺基亚手机，摩托罗拉手机，索爱手机。

3．限定和形容词

比如：

- 主机——免费主机，免费服务器。
- 电视机——电视机价格，电视机促销。
- 京东商城——京东商城官方网站。

上面提到的免费、促销、价格、官方网站这些附加限定词都很常见。便宜、怎样、是什么、好吗、评测、图片等也都是常见限定词。

3.5.7 网站流量分析

查看网站现有流量，用户都使用什么关键词搜索来到网站？经常能看到一些站长自己并没有想到的关键词。用户之所以能搜索这些关键词找到网站，说明搜索引擎认为你的网站与这种关键词有比较高的相关性。把这些关键词输入到 Google 工具，生成更多相关词，也是一个很好的关键词扩展方式。

3.5.8 单词交叉组合

上面提到的核心关键词、同义词、近义词、相关词、简写、地名品牌限定词等，放在一起又可以交叉组合出很多变化形式。如北京办公家具价格、京东商城电视机促销、诺基亚手机评价、云南旅游攻略等。

如果前面已经找到几百个关键词，交叉组合起来很容易生成数千个扩展关键词。这些比较长的组合起来的关键词可能搜索次数并不多，但数量庞大，累计起来能带来的流量潜力可观。

3.6 关键词分布

经过核心关键词确定与关键词扩展，应该已经得到一个至少包含几百个相关关键词的大列表。这些关键词需要合理分布在整个网站上。

优化多个关键词是很多初学 SEO 的人感到迷惑的问题。显然不可能把这么多关键词都放在首页上，否则页面内容撰写、链接建设、内部链接及锚文字的安排都将无所适从。

3.6.1 金字塔形结构

一个比较合理的整站关键词布局类似于金字塔形式。

核心关键词位于塔尖，只有两三个，使用首页优化。

次一级关键词相当于塔身部分，可能有几十个，放在一级分类（或频道、栏目）首页。意义最相关的两三个关键词放在一起，成为一个一级分类的目标关键词。

再次一级则放置于二级分类首页。同样，每个分类首页针对两三个关键词，整个网站在这一级的目标关键词将达到几百上千个。小型网站经常用不到二级分类。

更多的长尾关键词处于塔底，放在具体产品（或文章、新闻、帖子）页面。

3.6.2 关键词分组

得到关键词扩展列表后，重要的一步是将这些关键词有逻辑性地分组，每一组关键词针对一个分类。

举个例子。假设核心关键词确定为云南旅游，次级关键词可能包括昆明旅游、丽江旅游、大理旅游、西双版纳旅游、香格里拉旅游、临沧旅游等，这些词放在一级分类首页。

每个一级分类下，还可以再分二级。如大理旅游下又可以设置大理旅游景点、大理旅游地图、大理旅游攻略、大理美食、大理旅游交通、大理旅游自由行等，这些关键词放在二级分类首页。

再往下，凡属于大理地区内的景点介绍文章，则放在大理旅游景点二级分类下的文章页面。

这样，整个网站将形成一个很有逻辑的结构，不仅用户浏览起来方便，搜索引擎也能更好地理解各个分类与页面的内容关系。

有的行业并不像旅游一样有地区这种明显的划分标准，所以关键词分组的逻辑性并不直接明显。比如“减肥”这种词，就需要在进行关键词扩展时按行业常识将关键词分成多个组别。

经过关键词扩展得到大关键词列表后，按搜索次数排序，整体观察这些关键词可以从逻辑意义上分为几种，如图 3-15 所示。

减肥类关键词大致可以分为饮食减肥、减肥方法、局部减肥、快速减肥、运动减肥等一级分类。

	A	B	C	D	E
1	关键字	广告客户竞争程度	本地搜索量：12月	全球每月搜索量	
2	[减肥]	1	673000	673000	
3	[dhc 减肥]	0.4	673000	140	
4	[快速 减肥 方法]	0.93	450000	8100	
5	[瘦 身 减肥 方法]	0.66	450000	6600	
6	[运动 减肥 方法]	0.86	450000	1600	
7	[绿色 减肥]	0.66	301000	12100	
8	[减肥 公斤]	0.73	301000	-1	
9	[快速 减肥]	1	301000	165000	
10	[减肥 瘦 身]	1	301000	550000	
11	[减肥 法]	1	301000	8100	
12	[冬季 减肥]	0.86	301000	3600	
13	[减肥 汤]	0.86	301000	880	
14	[减肥 绝招]	0.73	301000	260	
15	[运动 减肥]	0.86	201000	14800	
16	[苹果 减肥]	0.93	201000	3600	
17	[节食 减肥]	0.86	165000	1300	
18	[醉谷 减肥]	0.6	165000	720	
19	[夏 日 减肥]	0.6	165000	260	
20	[腹部 快速 减肥 方法]	0.73	135000	2400	
21	[减肥 香蕉]	0.4	135000	28	
22	[饮食 减肥]	0.86	135000	1300	
23	[快速 减肥 瘦 身 方法]	0.8	135000	8100	
24	[苹果 餐 减肥]	0.6	135000	36	
25	[运动 减肥 瘦 身]	0.66	135000	390	
26	[运动 减肥 法]	0.73	135000	590	
27	[针灸 减肥 食谱]	0.66	90500	1600	
28	[减肥 餐]	0.86	90500	2400	
29	[快速 减肥 食谱]	0.8	74000	1000	
30	[腹部 减肥 法]	0.66	74000	320	
31	[减肥 塑身]	0.86	74000	14800	

Sheet1

图 3-15　相关关键词按搜索量排序

然后将所有关键词按上述分类分组，放在不同表内，如图 3-16 所示。

	A	B	C	D	E
1	[腹部 快速 减肥 方法]	0.73	135000	2400	
2	[腹部 减肥 法]	0.66	74000	320	
3	[脸 部 减肥]	0.73	60500	880	
4	[下半身 减肥]	0.8	22200	320	
5	[腹部 减肥 视频]	0.46	6600	390	
6	[腹部 减肥]	1	5400	33100	
7	[瑜伽 腹部 减肥 法]	0.4	480	320	
8	[大腿 减肥]	0.86	390	1900	
9	[腹部 减肥 的 最 好 方法]	0.66	390	1000	
10	[腰部 减肥]	0.86	390	1600	
11	[手臂 减肥]	0.73	390	720	
12	[腿 部 减肥]	0.8	260	1600	
13	[啤酒肚 减肥]	0.46	210	73	
14	[腿 减肥]	0.4	170	91	
15	[减肥 小腿]	0.53	170	110	

局部减肥 / 饮食减肥 / 快速减肥 / 健身减肥 / 总表

图 3-16　将关键词分组

可以看到，凡是和局部减肥有关的词就放在局部减肥表内，饮食减肥，快速减肥等有关词同样处理。

从一级分类列表里可以看出又能分为哪些二级分类，如局部减肥可以分为腹部减肥、大腿减肥、脸部减肥等。使用 Excel 将关键词合并、排序、分组后，整个网站的关键词金字塔结构就清晰地展现出来了。

3.6.3　关键词布局

进行关键词布局时还要注意以下几点。

（1）每个页面只针对两三个关键词，不能过多。这样才能在页面写作时有针对性，使页面主题突出。

（2）避免内部竞争。每个页面针对的两三个关键词，不要重复出现在网站的多个页面上。有的站长认为同一个词用多个页面优化，获得排名的机会多一点。其实这是误解，只能造成不必要的内部竞争。无论你为同一个关键词建造多少个页面，搜索引擎一般来说也只会挑出最相关的一个页面排在前面。使用多个页面反而分散了内部权重及锚文字效果，很可能使所有页面没有一个是突出的。

（3）关键词研究决定内容策划。从关键词布局可以看到，网站要策划、撰写哪些内容，在很大程度上是由关键词研究所决定的，每个版块都针对一组明确的关键词进行内容组织。关键词研究做得详细，内容策划就顺利。内容编辑部门可以依据关键词列表不停地制造内容，将网站做大、做强。虽然网站大小与特定关键词排名没有直接关系，但是内容越多，创造出的链接和排名机会就越多。

3.6.4　关键词-URL 对应表

关键词分组完成后，建议 SEO 部门将关键词搜索次数、目标 URL 等情况列表，如表 3-4 所示。

表 3-4　将关键词分配至 URL

关 键 词	百度指数	Google 月搜索	竞争指数	目标 URL	是否已收录	目前排名	目前月搜索流量
减肥瘦身	2057	301 000	10	www.domain.com	是	78	35
饮食减肥	121	135 000	8	www.domain.com/yinshi/	是	21	29
腹部减肥	247	5400	7	www.domain.com/jubu/fubu/	否	无	无

尤其重要的是，每一个重要关键词（网站首页及分类首页）都必须事先确定目标页面，不要让搜索引擎自己去挑选哪个页面与哪个关键词最相关。SEO 人员自己就要有意识地确定好每一个关键词的优化页面，这样在做内部链接时，凡是提到这个页面的关键词就链接到事先确定的 URL。

3.7　长尾关键词

长尾理论是著名的《连线》杂志主编 Chris Anderson 在 2004 年开始在《连线》杂志发表的系列文章以及后来出版的《长尾》这本书中具体阐述的。Chris Anderson 研究了亚马逊书店、Google 以及网上录像带出租网站 Netflix 等的消费数据，得出长尾理论。

3.7.1　长尾理论

所谓长尾理论，是指当商品储存、流通、展示的场地和渠道足够宽广，商品生产成本急剧下降以至于个人都可以进行生产，并且商品的销售成本急剧降低时，几乎任何以前看似需求极低的产品，只要有人卖，都会有人买。这些需求和销量不高的产品所占的市场份额总和，可以和少数主流产品的市场份额相比，甚至更大。

在传统媒体领域，大众每天接触的都是经过主流媒体（如电视台、电台、报纸）所挑选出来的产品，诸如各个电台每个月评选的十大畅销金曲，每个月票房最高的电影。图书市场也如此，权威的报纸杂志经常推出最畅销书名单。大众消费者无论自身品味差距有多大，在现实中都不得不处在主流媒体的狂轰滥炸之下，使得消费不得不趋向统一，所有的人都看相同的电影、书籍，听相同的音乐。

互联网及电子商务改变了这种情况。实体商店再大，也只能容下一万本左右的书籍。但亚马逊书店及 Netflix 这样的录像带出租网站，其销售场所完全不受物理空间限制。在亚马逊书店，网站本身只是一个巨大的数据库，网站能提供的书籍可以毫无困难地扩张到几万，几十万，几百万。

有各种各样奇怪爱好的消费者都可以在网上找到自己喜爱的书籍、唱片。网上书店可以出售非常另类、没有广泛需求的书，可以一年只卖出一本罕见的书给一个消费者，

营销成本并不显著增加。实体商店就无法做到这一点，不可能为了照顾那些有另类爱好的人，而特意把一年只卖一本的书放在店面里。实体商店货架展示成本是非常高的。

根据 Chris Anderson 对亚马逊书店、Netflix 网站及 Google 的研究，这种另类的、单个销售量极小的产品种类庞大，其销售总数并不少于流行排行榜中的热门产品。这类网站典型销售数字曲线如图 3-17 所示，也就是著名的长尾示意图。

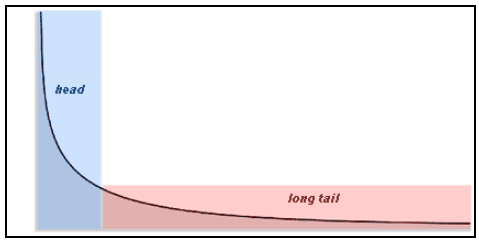


图 3-17　长尾效应示意图

横坐标是产品受欢迎程度，纵坐标是相应的销售数字。可以看到，最受欢迎的一部分产品，也就是左侧所谓的“头”，种类不多，单个销量都很大。“长尾”指的是右侧种类数量巨大，但单个产品需求和销售都很小的那部分。长尾可以延长到近乎无穷。虽然长尾部分每个产品销量不多，但因为长尾很长，总的销量及利润与头部可以媲美。这就是只有在互联网上才能实现的长尾效应。

3.7.2　搜索长尾

在 SEO 领域，较长的、比较具体的、搜索次数比较低的词就是长尾关键词。单个长尾词搜索次数小，但总体数量庞大，加起来的总搜索次数不比热门关键词搜索次数少，是流量金矿。另一个 SEO 人员关注长尾词的重要原因是，长尾词竞争小，使大规模排名提高成为可能。

搜索领域是长尾理论最明显的体现，因为渠道足够宽，每个网民的电脑都是渠道；送货、生产成本低，搜索引擎返回每一个关键词结果的成本几乎可以忽略不计；用户需求足够多元化，搜什么的都有。

搜索引擎工程师确认过，被搜索的关键词中有很大一部分搜索量很小，但总体数量庞大。甚至有的搜索词以前从来没有出现过，以后也再没有出现过，只被一个用户搜索过一次。

在长尾这个词被发明以前，SEO 行业早就确立了同样的关键词原则，只不过没有长尾关键词这个名词而已。长尾理论被提出以后，最先经常使用的就是 SEO 行业，因为这个词非常形象贴切地说明了大家一直以来已经在遵循的关键词选择原则。

很多站长从流量统计中也可以明显看到长尾现象。主要热门关键词就算排名不错，带来的流量也经常比不过数量庞大的长尾关键词。可以说，长尾关键词是大中型网站的

流量主力。大部分大型网站长尾流量应该占到一半以上，达到百分之七八十以上也不罕见。所以做好长尾是增加流量的关键之一。

长尾效应在小网站上较难发挥力量。大型网站主要关键词就算每天能带来几千访问量，与网站的几万甚至几十万日流量相比，还是个小小的零头。真正带来大量流量的，还是那几十万的长尾网页。小网站没有大量页面做基础，也无法有效吸引长尾搜索。

3.7.3 怎样做长尾关键词

做好长尾既简单又困难。说它简单是因为，一般来说不需要也无法做深入关键词研究，也不需要刻意优化特定长尾关键词。由于数量庞大，去查看搜索次数、专门调整页面优化都是不可行的，只能通过大量有效内容及网站结构方面的优化确保页面收录。只要页面基本优化做好，长尾关键词排名就能全面提高。

有的站长在论坛中询问，怎样做长尾关键词的研究？其实这很难研究。一个网站做几百几千个关键词的研究有可行性，但这个数量算不上是长尾词。真正体现长尾效应的网站至少要几万个页面，达到几十万数百万也只是普通的。长尾关键词数量至少上万。大致列出这些关键词，通过软件查询搜索次数或百度指数是可行的，但再进一步研究，如估计竞争程度、分配具体页面等，可行性及必要性大大降低。

所以做好长尾词的关键在于收录和页面基本优化，这两方面都是网站整体优化时必须要做的。不必考虑特定关键词，但结果是长尾词会全面上升。

说它难在于，做好长尾首先要有大量内容，对中小企业和个人站长来说，除了转载、采集，似乎没有更好的方法，除非网站是用户产生内容。

网站基本优化，尤其是内部链接结构，必须过关，才能保证大量包含长尾关键词的页面被收录。对一些大型网站来说，保证收录并不是一件简单的事。

要提高长尾词的排名，域名权重也是个因素。网上相同或相似内容很多，域名权重低，页面排名必然靠后。

所以长尾理论是 SEO 人员必须理解和关注的概念，但是要真正显示长尾关键词效果，却不能从关键词本身出发，而是从网站架构、内容及整体权重上着力。

3.8 三类关键词

按照搜索目的不同，关键词大致可以分为三种类型：导航类，交易类和信息类。

3.8.1 导航类关键词

导航类关键词指的是用户在搜索特定网站时，他知道自己想去哪个网站，只是不记

得网址或懒得自己输入网址，所以在搜索引擎直接输入品牌名称或与特定品牌有关的词。通常这类关键词排在第一的就应该是用户想访问的官方网站。

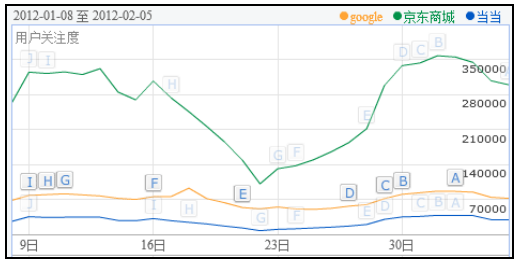


图 3-18　导航类关键词常常搜索量巨大

有的导航型关键词非常明确，比如 QQ 邮箱登录、京东商城官方网站。这种关键词最符合用户意图的结果通常只有一个，没有其他解释。有的导航型搜索稍微有些模糊，比如搜索京东商城、淘宝，用户既有可能是想访问京东商城或淘宝网站，也有可能是想看新闻或评价等。

如图 3-18 所示，导航类关键词常常搜索量巨大。导航类关键词占所有搜索的大致 10%，这是一个不小的比例。甚至在 Google 搜索百度（如图 3-19 所示）、在百度搜索 Google 的都大有人在。

关键字	广告客户竞争程度	本地搜索量：5月	全球每月搜索量	匹配类型：完全匹配
与所输入字词相关的关键字 - 按相关性排序				
[百度]		16,600,000	24,900,000	添加完全匹配
[百度网址]		数据不足	14,800	添加完全匹配
[百度网页]		数据不足	18,100	添加完全匹配
[搜狐 百度]		数据不足	46	添加完全匹配
[百 度 图片]		数据不足	590	添加完全匹配
[搜索 百 度]		数据不足	1,000	添加完全匹配
[badu百度]		数据不足	6,600	添加完全匹配
[baid百度]		数据不足	480	添加完全匹配
[daidu百度]		数据不足	8,100	添加完全匹配
[百度下载]		数据不足	12,100	添加完全匹配
[百度歌词]		数据不足	1,000	添加完全匹配
[百度网站]		数据不足	135,000	添加完全匹配
[百度音乐]		22,200	40,500	添加完全匹配

图 3-19　在 Google 搜索“百度”相关词

用户心里明知道想访问哪个网站，却在搜索引擎搜索，通常是因为现在的用户把搜索引擎当书签使用，懒得把网站放入收藏夹，也懒得自己输入网址，干脆到搜索引擎搜索，然后直接点击第一个结果。

自己的品牌名称被搜索时，网站排在第一是必需的。好在只要网站做得不是太烂，这相对容易做到。

竞争对手或其他相关品牌被当做导航型关键词搜索时对自己是个机会。在搜索广告领域，使用他人注册商标、品牌名称有一些争议。现在通常认为，用户搜索任何品牌、商标，任何人都可以对这个关键词竞价，但是在广告文案中不允许出现其他公司的注册商标和品牌名称。

在自然搜索方面限制更少，用户搜索竞争对手品牌时，你的网站排到前面并没有法律或道义方面的限制，只要不使用欺骗性手法，比如在页面上暗示与原商标、品牌有关系（其实没关系）、售卖劣质甚至非法产品。搜索导航类关键词，你的网站排在前面，从直接竞争对手的品牌搜索中获得流量，是一个可以接受而且目标比较精准的方法。

当然，在自己的页面中出现和突出竞争对手品牌名称一般不太适合。要想在竞争对手的导航型关键词中排名靠前，增强外部链接和锚文字是最主要的方法。

3.8.2 交易类关键词

交易类关键词指的是用户明显带有购买意图的搜索。比如“电视机网上购买”、“索爱手机价格”等。交易型关键词占全部搜索的 10%左右。

显然交易类关键词的商业价值最大，用户已经完成商品研究比较过程，在寻找合适的卖家，离网上交易只有一步之遥。吸引到这样的搜索用户，转化率是最高的。所以在进行关键词研究时，发现这类交易意图比较明显的关键词，优先度应该放在最高，可以考虑特殊页面专门优化。交易类关键词在网站上的分布需要非常精确，把用户直接导向最能说服用户购买的页面，而不是分类或帮助等无关页面。

3.8.3 信息类关键词

信息类关键词指的是没有明显购买意图，也不含有明确网站指向性的搜索。如“手机图片”、“减肥方法”等。这类关键词占总搜索数量的其他 80%。

信息类搜索数量最多，变化形式也最多。用户通常还处在了解需求、商品研究阶段。针对信息类关键词进行优化，是网站长尾页面的任务。虽然这类关键词并不一定导致购买，但是在用户进行商品研究时进入用户视野也是非常重要的。好的网站设计、出色的文案，让搜索信息的用户记住网站或品牌名称，经常会导致用户以后直接搜索网站名称，也就是导航类关键词，进而实现转化。网站内容越多，出现在信息类关键词结果的机率越大。

3.9 预估流量及价值

正规公司关键词研究的最后一步是预估搜索流量及价值。

个人站长做关键词研究不一定需要这一步。找到最合适的关键词就可以直接去做，尽力能做多少就做多少。但正规公司，尤其是大公司则不行。整个 SEO 项目是否能获得批准，能否获得公司高层支持，申请预算、安排人员、工作流程及时间表等，都需要 SEO 人员提供明确的预计搜索流量及给公司带来的价值。

3.9.1 确定目标排名

要预估搜索流量，首先需要根据前面得到的关键词竞争指数及公司本身的人员、资金投入，预计网站关键词可以获得什么样的排名。

前面做了关键词研究的所有核心关键词及扩展关键词都应该有预计排名位置。当然，这种预计与个人经验、团队决心有很大关系，与最后的实际结果不可能完全吻合，完全符合预计只是巧合而已。这也是预估流量不可能很精确的第一个原因。

由于种种原因，项目执行下来能够达到预计排名的比例不可能是百分之百，只可能实现其中一部分关键词排名。所以在预估流量时不能按照所有目标关键词都达到预期排名计算，只要能完成 30%～50%就已经不错了。好在网站通常还会获得一些没有研究过的关键词排名，使误差减小。

3.9.2 预估流量

现在 SEO 手上已经有了三组数据：关键词搜索次数，关键词预计排名，搜索结果页面各排名位置的点击率。有了上面三组数据，再考虑到百度和 Google 搜索市场占有率，就可以计算出各关键词及整个网站预计能得到的搜索流量，如表 3-5 所示。

表 3-5　预估流量示例

关 键 词	Google 月搜索次数	预计排名	点 击 率	预计 Google 流量
减肥	673 000	9	2.8%	18844
运动减肥	201 000	8	3.0%	6030
快速减肥	301 000	8	3.0%	9030
饮食减肥	135 000	5	4.9%	6615
节食减肥	165 000	5	4.9%	8050
苹果减肥	8100	3	8.5%	688
腹部减肥	5400	8	3.0%	162
脸部减肥	60 500	7	3.4%	2057
快走减肥	1000	5	4.9%	49
跑步减肥	1300	5	4.9%	64
预计 Google 月搜索流量				57786
Google 市场份额				33.2%
预计总搜索流量/每月				174 054
排名达成 50%时月搜索流量				87 027
排名达成 30%时月搜索流量				52 216

表 3-5 所示只是示例。读者做自己网站关键词研究时应该不止列出 10 个关键词。另外，考虑到无法预见的长尾关键词，总搜索流量还会增加。

这样预估出来的流量不可能非常精确，除了预计排名的不准确、百度和 Google 排名也不可能相同外，还有更多方面的原因。

1．搜索次数

首先，关键词搜索次数很可能就不准。Google 关键词工具显示的数字虽然明确标为就是搜索次数，但很多人的统计数据表明，这个次数并不准确。一般 Google 给出的次数要高于实际次数。

而且实际有效搜索次数与发生过的搜索次数也有落差。比如搜索“SEO”或“搜索引擎优化”的查询次数可能不低，但其中有很大一部分是 SEO 人员在查看排名而已，不会点击其中任何一个结果，这些是无效搜索。

一个矫正搜索次数的方法是，找出自己网站上（或能拿到流量数据的其他网站）已经有不错排名的关键词，列出现有排名位置及对应的真实搜索流量，就可以计算出实际搜索次数与 Google 关键词工具给出的搜索次数之比，如表 3-6 所示。

表 3-6　根据实际流量估计较准确搜索量

关 键 词	搜索引擎优化
排名	2
月实际 Google 流量	174
第 2 位点击率	11.9%
实际有效搜索次数	1462
Google 工具显示搜索次数	8100

像前面提到的，搜索“搜索引擎优化”的很多是在查排名的 SEO 人员，不会点击任何页面，所以实际有效搜索次数与 Google 关键词工具显示的搜索次数相差巨大。普通关键词没有这么大差别，但通常 Google 显示的数字还是多出 30%～50%。

同样的原理也适用于百度指数。百度官方并没有说百度指数就是搜索次数，而只是关键词的受关注程度。有的 SEO 人员认为百度指数就是搜索次数，有的人认为这两者相差很大。如果你知道某个关键词的排名及实际流量，就可以计算出百度指数与百度实际搜索次数的比例关系。

需要注意的是，这个比例并不是恒定的。不同行业关键词实际搜索次数与 Google 关键词工具显示搜索次数或百度指数之比并不相同，这可能与不同目标市场的用户行为方式不同有关。所以 SEO 人员需要针对自己的网站及所在行业自行计算比例，不能照搬别人的数据。

当然，样本数越多，数据越准确。只计算一个关键词不能准确说明情况。

2．点击率

搜索结果页面各排名位置点击率也不精确。我们在第 2 章介绍的搜索排名点击率来自于美国在线 AOL 泄露的数据，与 Google 点击率有一定差别，更不完全适用于中文搜索页面，特别是百度。

2010 年 5 月，网络广告公司 Chitika 又发布一份搜索结果点击率报告，如表 3-7 所示。

表 3-7　Chitika 发布的搜索结果点击分布

Google 结果	显示次数	点 击 率
1	2 834 806	34.35%
2	1 399 502	16.96%
3	942 706	11.42%
4	638 106	7.73%
5	510 721	6.19%
6	416 887	5.05%
7	331 500	4.02%
8	286 118	3.47%
9	235 197	2.85%
10	223 320	2.71%
11	91 978	1.11%
12	69 778	0.85%
13	57 952	0.70%
14	46 822	0.57%
15	39 635	0.48%
16	32 168	0.39%
17	26 933	0.33%
18	23 131	0.28%
19	22 027	0.27%
20	23 953	0.29%

显示为曲线更为直观，如图 3-20 所示。

Chitika 统计得到的数字与 AOL 泄露数据得到的点击率数字有比较大的差别。到底哪个数据更反映实际情况，整个 SEO 行业没有定论，不同行业、不同网站很可能会得到不同数据。

因此，有条件的团队应该尽量多收集已知关键词排名的排名位置和真实搜索流量数据，计算出比较可靠的搜索结果点击率，样本越多，结果准确度也越高。有些公司使用软件计算海量关键词排名数据，对预估流量有很大意义。

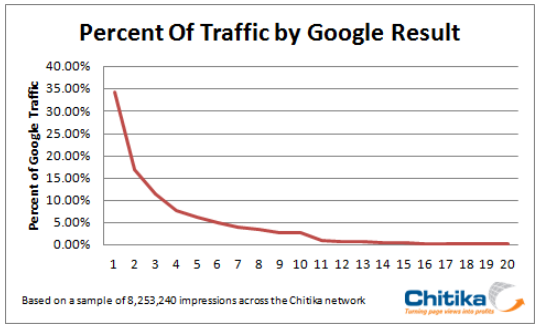


图 3-20　Chitika 发布的搜索结果点击分布曲线

除了上面提到的两个主要原因，预估流量不准确还可能出于下面几个原因。

1．整合搜索

搜索结果中出现的视频、图片、新闻等内容对用户视线及点击的影响很大，使整个页面各个结果的点击率出现波动。

2．多次点击

用户每次搜索可能产生不止一次点击，而是多次点击几个结果。而每次搜索到底产生多少次点击，与行业及搜索引擎对这个关键词排名的质量高低有很大关系，并没有固定的数字。

3．关键词类型

不同类型关键词搜索点击情况也不同。比如，一般来说，导航型搜索第一个结果的点击率比其他类型关键词要高得多。用户知道应该访问哪个网站，只是把搜索引擎当做方便的书签，根本不会去看第二名及以下的结果。

4．PPC 影响自然排名点击

有人做过统计，右侧或上部出现竞价广告时，对左侧自然排名中同一个公司的结果点击率有帮助，使点击率提高。这可能与用户认为有实力做广告的公司更可靠有关。

5．季节性波动

下一节要讨论的关键词搜索次数季节性波动也会造成预估流量的不准确。如果查询记录的搜索次数刚好是在搜索高峰期，得到的预估流量可能比实际的高出几倍甚至几十倍。

无论如何，就算预估流量不可能非常准确，甚至可能与实际情况相差几倍之多，但重要的是 SEO 人员能知道未来预期流量在哪个数量级，建立心理预期，并为公司高层提供做判断、决策时需要的数字。“预估流量是 5000，实际结果是 2000”是可以接受的。但是如果项目开始时，连最终流量会是几百或几千或几万都毫无概念，说服高层批准项目和预算将变得很困难。

3.9.3 预估搜索流量价值

流量并不是目标，订单和赢利才是目标。得出预计流量后，结合网站转化率及平均订单销售额和平均每单毛利，就可以计算出预期通过 SEO 获得的搜索流量能带给公司多少实际价值。

如预计搜索流量为每月 8 万个独立 IP，网站历史平均转化率为 1%，则搜索流量将带来 800 个订单。如果平均每单金额是 100 元，毛利 30 元，则搜索流量每月将贡献 8 万元销售额，2 万 4 千元毛利。

上面的计算已经很简化。搜索流量转化率不一定等于网站平均转化率。有的网站搜索流量转化率远高于平均转化率，因为搜索流量质量较高，用户有较强的购买意图。也有的网站搜索流量转化率低于平均转化率，因为回头客很多，直接点击流量转化率很高。所以如果网站有一定的累积数据，还要根据流量分析数据做矫正。

有的网站不直接销售产品，则需要根据网站目标计算出每次转化的价值，再计算搜索流量的价值。后面效果监测一章再深入探讨。

3.10 关键词趋势波动和预测

前面研究关键词搜索次数时，通常只看一段时间的搜索次数，比如一个月之内。但绝大多数关键词搜索次数会随时间波动，做关键词研究时，除了当前搜索次数，还要考虑随时间变化情况。

3.10.1 长期趋势

一部分关键词随时间具有稳定上升或下降趋势，如图 3-21 所示为 Google 趋势。

同样是电子产品，录像机和 DVD 机都已经是明日黄花，关注的人越来越少，而智能手机越来越受关注。这种关键词搜索长期趋势对网站主题及内容的选择，甚至产品研发、是否进入某个行业都有决定性影响。搜索次数持续下降的关键词不一定就不能做，但企业和 SEO 人员都应该有正确的心理预期，就算排名很好，流量还是会下降。

图 3-21　关键词搜索量长期趋势

3.10.2　季节性波动

有很多关键词随季节正常波动。最明显的是各个节日，在节日周围一段时间搜索量剧增，而其他时间很少有人关心。

与特定节日或时间相关联的产品信息也随之产生季节性波动，如粽子、鲜花、巧克力、短语、节日祝福、高考等关键词，如图 3-22 所示。

图 3-22　关键词搜索量周期性波动

有的关键词季节性波动不太容易从表面想象得到，如图 3-23 所示 Google 趋势显示的两个词。

关键词“液晶电视”一年之中有三次搜索高峰，分别在圣诞新年、四月初、十一之前。十一和新年期间应该是正常购物高峰，但四月初为什么会出现购买液晶电视的高峰？有点令人费解，也许对行内人来说是很简单的原因。可以肯定的是，每年搜索液晶电视的次数都是这样的规律，不像是某个产品或型号推出造成的效应。

而搜索“家具”的次数每年春节期间急剧下降。一般来说，春节期间上网的人数本身就会大幅减少。不过从与搜索液晶电视的次数相比可以看到，整体上网人数的下降并不能完全说明为什么搜索家具的在春节期间会下降这么多。

图 3-23　关键词搜索量波动不一定显而易见

这种随季节波动比较大的关键词，SEO 人员应该事先了解趋势，提前做出内容建设、外链建设等方面的准备，有时候可能还需要开设专题，进行产品促销等。由于搜索引擎收录页面、计算排名都需要一段时间，因此 SEO 人员针对这些季节性关键词做排名也必须提前准备。

3.10.3　社会热点预测

每一次出现社会热点新闻，都会带动一批关键词搜索次数大增。

笔者在自己的博客上做过一个简单的实验。2010 年 1 月 13 日，Google 第一次宣布退出中国，我早上看到这个新闻后，意识到这是比较少出现的、与 SEO 还算有点关系的有可能成为社会热点的新闻，所以赶紧在博客上发了一篇帖子，标题就是：Google 退出中国，如图 3-24 所示。

图 3-24　“Google 退出中国”搜索量变化

帖子内容其实非常简单，只是列出了我看到的 9 篇其他人的相关博客帖子，如此而已，没有自己的评论。Google 抓取和索引速度非常快，当天我的这篇帖子在搜索“Google 退出中国”这个词时排到了第五位左右，如图 3-25 所示。

从 GA 统计可以看到，13 日、14 日、15 日这几天博客流量大增，14 日独立 IP 数是 7000 多，其中绝大多数是因为这个热点新闻搜索而来的，如图 3-26 所示。

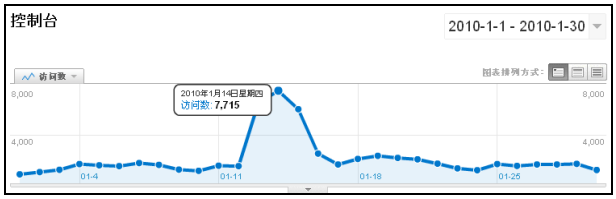


图 3-25　因“Google 退出中国”这个词流量大增

2010年 1 月 11	3615	14195	58541	1.20 G字节
2010年 1 月 12	3723	10359	53510	815.21 M字节
2010年 1 月 13	10338	19592	130245	1.70 G字节
2010年 1 月 14	12178	23095	150771	2.11 G字节
2010年 1 月 15	9689	18981	121768	1.87 G字节
2010年 1 月 16	4348	10182	60120	937.56 M字节
2010年 1 月 17	3262	8304	44466	697.72 M字节

图 3-26　AWStats 显示的流量

其实查看一下 AWStats 统计数字可以看出，实际 IP 数已经上万，但是因为我的博客只是放在虚拟主机上，那两天的流量已经把主机拖得非常慢，大部分用户打开网页比较困难，GA 的 JS 代码很多时候没有被运行，流量也没有被 GA 记录。

从这个简单的案例可以看出，社会热点新闻搜索次数非常高。我的博客帖子只是在 Google 排在第五六位，如果能排在第一位，流量应该翻上几倍。如果能进入百度前几位，流量又会翻几倍。如果是更能引起普通网民注意的社会热点，搜索次数又不知会翻多少倍。

当前的中国是一个绝不缺乏社会热点新闻的时代。这个门、那个门几乎每星期都会出现，网上热点层出不穷。有很多站长就是通过捕捉社会热点关键词，给网站带来每天几万几十万的独立 IP 流量。

捕捉社会热点最好的工具当然是百度风云榜和谷歌热榜（可惜目前谷歌热榜已关闭），在 SEO 工具部分还有更详细的介绍。这些关键词排行榜除了列出稳定的热门词，还列出上升最快的关键词，这些关键词就是最有潜力成为热点的词。很多专门做社会热点的 SEO 人员每天都会留意上升最快的关键词，看到合适的，立即组织内容甚至专题。

对社会新闻的敏感度是捕捉热点关键词的关键。除了关键词排行榜，电视报纸上的新闻都可能是灵感的来源。谁能先捕捉、预计到哪件事会成为热点，谁就将得到这些流量。

除了针对热点组织相关内容，拥有一个权重比较高、爬行频率也高的域名也有很大帮助，否则有了内容却不能被快速收录，也无法达到效果。

当然，通过热点关键词带来搜索流量，不一定适合所有网站。前面讨论过，有效流量才是我们需要的。一个电子商务网站通过无关的社会热点带来流量，价值到底有多大，要看网站营销人员是否能把自己的产品与新闻建立一定的逻辑或情感联系。如果从哪个角度看都毫无关系，这样的流量就没有什么意义了，除非网站纯粹靠广告赢利。

3.11 竞争对手研究

竞争对手研究是市场竞争研究的重要部分，对判断特定关键词竞争程度及了解行业整体情况非常有价值。

确定 SEO 方面的竞争对手很简单，在搜索引擎搜索核心关键词，排在前十名到二十名的就是你的主要竞争对手。SEO 人员需要从下面几方面了解竞争对手情况。

3.11.1 域名权重相关数据

域名权重在很大程度上决定了整个网站的排名能力。需要查看的数据如下。

1．域名年龄

既包括域名最初注册时间，也包括网站第一次被搜索引擎收录的时间。不过最早被搜索引擎收录的时间没有办法直接查到，通常可以借鉴互联网档案馆上网站最早出现的日期：

http://www.archive.org/web/web.php

如图 3-27 所示，我的博客所在域名第一次出现内容是 2005 年 12 月，与实际情况基本符合。点击每个日期链接，还可以看到当时网站页面内容。

域名越老，权重越高。如果你有一个 10 年的老域名，针对中等竞争程度的关键词，你将有很大优势。如果你拥有一个一九九几年就注册了的域名，不用看其他部分，域名本身就是个宝藏。当然如果竞争对手是这么老的网站，超越难度也很大。

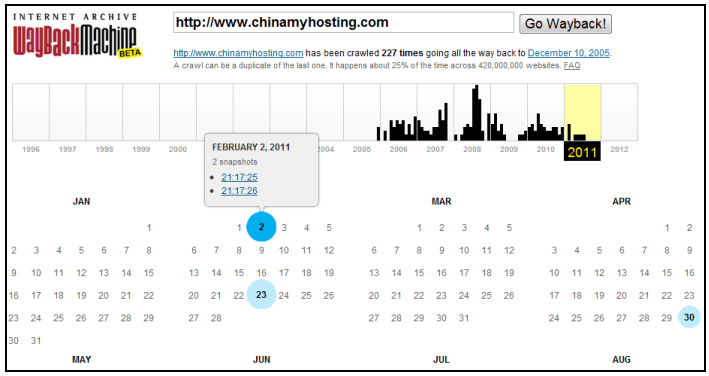


图 3-27　互联网档案馆记录网站历史

2．Google PR 值

虽然 PR 值在 Google 排名中的重要性近来下降很多，而且 Google PR 值也不能说明网站在其他搜索引擎的权重，但 PR 值毕竟是直观说明外部链接质量和数量的最好指标。一般来说首页 PR 值达到五或六，是一个不错的网站。达到七八及以上，权重就很高了。

3．快照日期

快照越新，说明搜索引擎爬行收录的频率越高，这也在某种程度上代表了域名权重。

4．收录页面总数

一般来说网站越大，权重越高，收录数越多，同时说明网站整体结构比较合理。

5．外部链接情况

包括域名总链接数、首页外部链接数量、内页外部链接数量、外部链接总域名数、锚文字分布等。还可以进一步查看外部链接来自于哪些网站，是否以交换链接、论坛、博客留言等低质量链接为主。

6．主要目录收录情况

如雅虎目录、开放目录、Hao123 等。这些高质量的目录只收录质量、权重比较高的网站，搜索引擎也清楚这一点，所以对被收录的网站会给予相应的权重提升。

7．社会化媒体出现情况

主要包括网上书签服务收录数目，如 Delicious、百度搜藏等。在社会化内容推荐网络出现的次数，如 Digg、Twitter 等网站被推荐或被提到的次数。在社会化媒体中出现次数多，说明网站受用户欢迎，搜索引擎也很可能把这一点计入排名算法中。

上面这些数据如果一个个人工查询，耗时耗力。幸运的是有工具可以帮助完成。比较方便的工具包括 SEO 工具章节将会详细介绍的 SEO for Firefox 和 SEOQuake。

SEO for Firefox 界面简单干净，使用方便，可惜不支持百度，如图 3-28 所示。

图 3-28　SEO for Firefox 显示页面基本 SEO 信息

SEOQuake 支持百度，如图 3-29 所示。

通过这两个工具查询特定关键词，与权重有关的网站数据会清楚地标明在搜索结果页面上，各个竞争对手网站权重如何一目了然。

图 3-29　SEOQuake 显示百度搜索结果的页面 SEO 信息

3.11.2　网站优化情况

研究竞争对手实力和关键词难度时，对手网站本身优化情况也要考虑在内，包括网站结构及页面针对关键词的优化。

有时排在前面的页面仅仅是因为外部链接比较强，主要竞争对手优化得都不是很好，这就存在更大的市场机会。

什么样的网站优化是合理的？这是本书后面网站结构及页面优化两章要深入阐述的内容，读者看完本书就会很清楚。在动手优化网站之前，通常需要深入分析优化最好的一两个竞争对手网站，仔细研究每一个细节，常常会学到很多新的有效的优化技巧。

在做关键词研究时如果不能非常详细地分析所有对手，至少要看以下几个要点：

- 页面标题标签是否包含关键词？
- 网站栏目分布是否清晰合理？
- URL 是否静态化？
- 网站链接结构是否合理有效？
- 内页距离首页点击距离有多远？是否能在三四次点击内到达所有内页？
- 网站主要页面是否有实质内容？
- 导航系统是否使用了 Flash，JS 脚本等不利爬行的方法？
- 页面是否使用了 H1 和 H2 标签？是否在其中包含了页面目标关键词？

经过优化的网站在各个方面都能看出痕迹。而没有优化的网站，通常连上面这些最基本的地方都会有问题。所以简单查看几处重要元素，就能看出网站是否经过了基本的 SEO。

3.11.3 网站流量

了解竞争对手网站流量，既能更确切地知道对方排名实力，也能在一定程度上印证预估流量数字。

竞争对手网站真实流量外人无法得到确切数字，除非有卧底。SEO 工具部分详细介绍的 Google 趋势和 Google AdPlanner 可以辅助查询竞争对手网站大致流量。

Google 趋势统计的流量数据不仅仅来自于 Google 提供的免费流量分析工具 Google Analytics，还有 Google 工具条、关键词排名点击率、网络服务商 ISP 等来源。就笔者了解的真实流量的一些网站看，Google 趋势显示的流量大部分情况下基本准确。流量绝对数字误差肯定存在，有时候能相差 1 倍，但这已经能满足需要，比瞎猜强多了，至少不会出现数量级的误判。

与流量绝对数字比，Google 趋势显示的流量变化趋势则更为准确，如流量高峰和低谷出现的时间点，流量增长或降低的长期趋势，与实际情况相当符合，如图 3-30 所示。

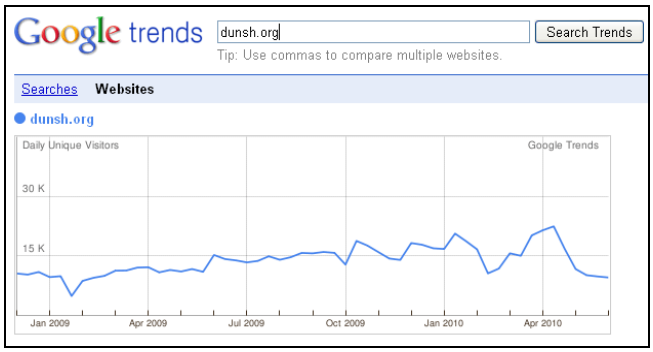


图 3-30　Google Trends for Websites 显示网站流量

3.12 快速网站诊断

详细诊断网站、提出优化方案是本书要解决的问题，本节先简单介绍网站快速诊断方法。SEO 服务商在刚刚与客户接触时，尤其需要对目标网站做快速检查，发现重要问题。

快速诊断的第一步是检查上一节提到过的、研究竞争对手网站时同样的指标。唯一要加上的是计算页面收录比例，也就是搜索引擎收录页面数与网站实际总页面数之比。站长一定要清楚自己网站上的实际页面总数，如果收录比例在 20%～30%，说明网站结构有很大问题，需要改进。达到 60%以上才是合格的。

快速诊断的第二步是查看 Google 网站管理员工具中给出的信息。

http://www.google.com/intl/zh-CN/webmasters/

Google 网管工具控制台如图 3-31 所示。

图 3-31　Google 网管工具控制台

Google 网管工具是一个非常好用、非常强大的工具，而且免费，强烈建议所有站长注册。微软必应、雅虎也都有类似工具，不过提供的数据与 Google 网管工具相比过于简单。百度目前也在内测网管工具。

首先，如果网站有重大问题，如网站被黑、页面有病毒或者网站因为作弊被删除和严重惩罚，Google 会在网管工具给站长留言，站长不必猜测到底出了什么问题。

没有严重问题时，站长可以查看下面这些主要内容。

3.12.1　robots 文件检查

整个网站不能收录或某个目录下所有页面都不能收录，经常是因为 robots.txt 文件差错引起的。网管工具抓取工具权限部分显示出 Google 所抓取的 robots 文件内容，如图 3-32 所示。

站长也可以在这里试验不同的 robots 文件指令，然后输入一个网址，测试网址是否可以被收录，或是被禁止，如图 3-33 所示。

robots 文件中的任何一个字母差错都可能造成致命影响。有了这个工具，站长可以确保 robots 文件中的每一行代码正确，不会错误禁止应该被收录的文件或目录。

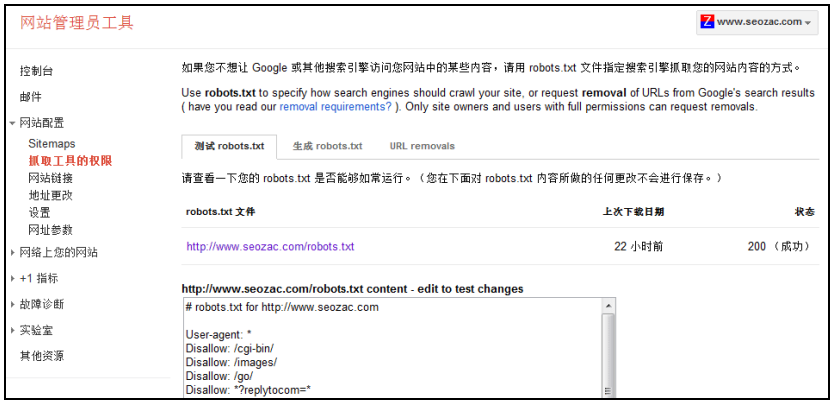


图 3-32　Google 网管工具显示 robots 文件内容

图 3-33　在 Google 网管工具测试页面是否被 robots 文件禁止爬行

3.12.2　首选域设置

站长可以设置 Google 应该收录带 www 还是不带 www 的网址版本，称为首选域，如图 3-34 所示。

图 3-34　Google 网管工具设置首选域

当然，在 Google 网管工具设置的首选域对百度等其他搜索引擎完全不起作用。这只

是解决 Google 网址规范化的辅助手段，不能完全依靠这个设置，正确合理的网站结构才是解决问题的根本方法。站长也可以在这部分设置网站目标地理区域。

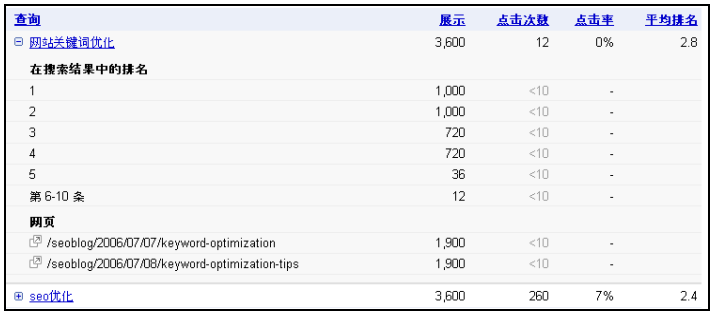
3.12.3 关键词排名

在搜索查询部分，网管工具列出网站获得排名的关键词有哪些，并且列出了搜索结果显示次数、点击次数、点击率和平均排名，如图 3-35 所示。

图 3-35　Google 网管工具显示的关键词排名

在搜索引擎工作原理章节介绍了搜索结果页面前十名结果的点击率，网管工具则列出了网站真实排名及点击数字。这也为 SEO 人员提供了搜索结果点击分布的另一组数据，可以用于搜索流量预估。不过，要注意的是，网管工具中列出的点击率，很多时候与网站权重、知名度、页面标题标签的写作有很大关系，并不一定符合其他关键词的点击情况。

比如笔者的博客在搜索“网站关键词优化”时，显示在第一位时为 1000 次，点击率还不到百分之一，如图 3-36 所示。



查询	展示	点击次数	点击率	平均排名
网站关键词优化	3,600	12	0%	2.8
在搜索结果中的排名				
1	1,000	<10	-	
2	1,000	<10	-	
3	720	<10	-	
4	720	<10	-	
5	36	<10	-	
第 6-10 条	12	<10	-	
网页				
/seoblog/2006/07/07/keyword-optimization	1,900	<10	-	
/seoblog/2006/07/08/keyword-optimization-tips	1,900	<10	-	
seo优化	3,600	260	7%	2.4

图 3-36　点击率很低的关键词

这与一般搜索结果点击分布相差甚远，说明得到排名的博客帖子与用户搜索意图有一定距离，也可能是因为标题写得不好，造成点击率如此之低。

另外一个关键词“SEO 优化”，排在第二位时点击率是 9%，排在第三位时为 7%，大致正常，如图 3-37 所示。

查询	展示	点击次数	点击率	平均排名
⊞ 网站关键词优化	3,600	12	0%	2.8
⊟ seo优化	3,600	260	7%	2.4
在搜索结果中的排名				
1	<10	<10	-	
2	2,900	260	9%	
3	170	12	7%	
4	<10	<10	-	
第 6-10 条	<10	<10	-	
第 2 页	91	<10	-	
网页				
/seoblog	3,600	260	7%	
⊞ seo是什么	2,900	260	9%	3.4

图 3-37　点击率正常的关键词

而另一个搜索词“Zac”，点击率又走向另一个极端，排在第 6～10 位时点击率高达 23%，远远高出一般搜索点击分布，如图 3-38 所示。

查询	展示	点击次数	点击率	平均排名
⊞ 网站关键词优化	3,600	12	0%	2.8
⊞ seo优化	3,600	260	7%	2.4
⊞ seo是什么	2,900	260	9%	3.4
⊟ zac	1,600	210	13%	10
在搜索结果中的排名				
1	28	<10	-	
2	390	73	19%	
3	260	<10	-	
4	46	<10	-	
第 6-10 条	320	73	23%	
第 2 页	590	73	12%	
第 3 页及后续页面	22	<10	-	
网页				
/seoblog	1,000	210	21%	
/seoblog/about	590	28	5%	
/seoblog/2008/12/18/google-serp-test	<10	<10	-	
/seoblog/2009/04/24/smx-xiamen-pics	<10	<10	-	
（索引页）	<10	<10	-	
⊞ 搜索引擎优化	1,300	110	8%	2.2

图 3-38　点击率很高的关键词

这说明搜索“Zac”时，有很大一部分用户确实是在寻找笔者的博客。

3.12.4　外部链接

Google 的 link：指令非常不准确，基本不能用来查看外部链接。网管工具中列出的外部链接则要准确得多，SEO 人员可以一目了然地看到自己网站上哪些页面最受欢迎，吸引到最多的外部链接，如图 3-39 所示。

概述 » 所有链接的网页
链接自其他域的网页。

下载此表　下载更多示例链接　显示 10 行　第 1 到 10 条记录（共 560 条）

您的网页	链接数量	来源域
http://www.seozac.com/	129,665	1,954
/interviews/interview-matt-cutts-en/	1,863	46
/interviews/interview-matt-cutts-cn/	1,842	39
/marketing/	820	125
/onpage/title-tag/	477	129
/google/google-pr/	471	111
/content/seo-biggest-factor/	449	103
/seo/301-redirect/	433	121
/google/	374	91
/about/	374	83

第 1 到 10 条记录（共 560 条）

图 3-39　Google 网管工具显示的外部链接

Google 网管工具还列出了外部链接使用的锚文字，网管工具里称为“定位文字”，如图 3-40 所示。

概述 » 定位文字

定位文字	
1. reply	101. yahoo 排名技术研究
2. zac seo	102. april 2009
3. 搜索引擎优化 seo 每天一贴	103. january 2009
4. seo 每天一贴	104. 用户体验和行为方式
5. zac 的 seo 博客	105. 外部链接策略
6. http www seozac com	106. august 2010
7. www seozac com	107. february 2010
8. zac	108. march 2009
9. seo 每天一贴 zac 博客 《 seo 实战密码 》已出版 正翻译另一本	109. june 2006
10. zac seo 每天一贴	110. november 2008
11. seozac	111. october 2007
12. seo 书	112. august 2009
13. in chinese	113. december 2006
14. interviewed google's matt cutts	114. seo 问题解答
15. seo 每天一帖	115. 英文网站 seo
16. 搜索引擎优化每日一贴	116. seo 优化论点
17. http www chinamyhosting com seoblog	117. 其他搜索引擎排名
18. seo 每日一帖	118. december 2009

图 3-40　Google 网管工具显示的外部链接锚文字

这也使 SEO 人员可以清楚地看到自己想排名的核心关键词是否有足够的外部链接锚文字支持。

可惜网管工具外部链接不能用于查询其他网站，只能看到站长自己认证过的网站数据。

3.12.5　网站内容

网管工具关键词部分实际上列出的是 Google 在网站上抓取的最常见关键词，如图 3-41 所示。显然这些最常见的关键词就反映了网站的内容主题。

图 3-41　Google 看到的网站最常见关键词

从抓图中可以看到，笔者的博客主体内容大致合理，但是最常出现的两个词却出人意料，一个是 Reply（回复），一个是 says（说）。这两个英文词是博客评论模板中出现的，每一个评论都会出现 Reply 和 says 各一次。当评论条数比较多时，整个网站主体内容产生了一定的偏移，如图 3-42 所示。

图 3-42　博客评论中大量出现的“Reply”和“says”

查看这里列出的关键词，对页面尤其是首页的文案撰写和修改有重要意义。

3.12.6　内部链接

内部链接部分列出所有页面的内部链接数，如图 3-43 所示。

图 3-43　Google 网管工具显示的内部链接

站长从这里可以大致判断网站内部链接结构是否有重大缺陷。如果全站主导航中出现的分类首页内部链接数非常低，很可能说明导航系统有问题。

内部链接数的另外一个作用是反映出网站收录页面数。Google 的 site：指令也不太准确，而且现在越来越不准确，经常不能反映出收录数字。网管工具中内部链接部分列出的首页内部链接总数，大致上就相当于 Google 收录的页面总数，因为网站上每一个页面都应该有到首页的链接。

3.12.7　抓取错误及统计

抓取错误部分列出 404 错误（页面不存在）、被 robots 文件禁止而不能收录的页面等，如图 3-44 所示。

图 3-44　Google 网管工具显示的 404 错误

其中 404 错误对检查网站上是否存在错误链接很有用。对每个 404 错误，网管工具都列出了到这个网址的链接，如图 3-45 所示。

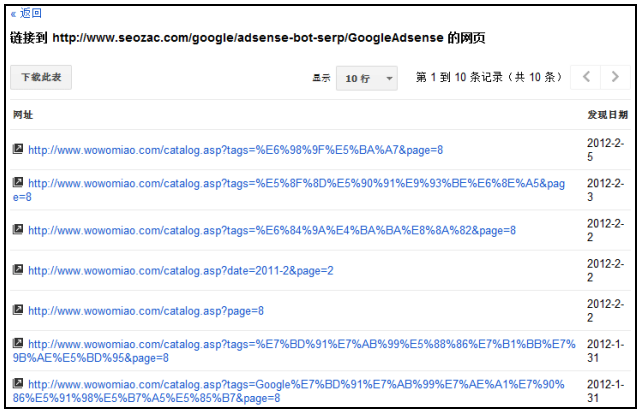


图 3-45　指向不存在页面的链接

如果到不存在页面的链接是网站内部发出的，说明这些链接页面上的链接地址有错误。如果链接向不存在页面的是其他网站，站长可以尝试联系对方更改错误链接到正确位置。

网管工具也列出抓取统计信息。

如图 3-46 所示，笔者的博客每天抓取页面数平均为 400 多页，对于一个只有几百篇帖子的博客来说很正常。如果是几十万上百万页面的大中型网站，平均每天抓取几百个页面就太糟糕了，照这个速度，要把整个网站更新一遍，恐怕得好几年时间。大中型网站每天抓取几万页面是最低标准。如果抓取频率太低，SEO 人员必须在外部链接建设及网站结构上多下工夫。

图 3-46　Google 网管工具显示的抓取统计信息

3.12.8　HTML 建议

查看 Google 的 HTML 建议是寻找网站上可能出现的复制内容的最简便方法，如图 3-47 所示。

网管工具列出了重复的元说明（也就是说明标签）及标题标签个数和具体页面。笔者的博客存在的重复标题标签如图 3-48 所示。

标题标签重复，实际上经常意味着这些页面本身内容重复，往往是网站结构造成的。要注意的是，有时候网管工具中列出的数据并不完整，一般来说博客上标题标签重复的页面绝不止两页。

图 3-47　Google 网管工具显示的 HTML 建议

图 3-48　Google 网管工具显示的重复标题标签

3.12.9　模拟蜘蛛抓取

站长可以输入自己网站上的任何一个网址，网管工具会发出 Google 蜘蛛，实时抓取页面内容，并显示出抓取的 HTML 代码，包括服务器头信息和页面代码，如图 3-49、图 3-50 所示。

图 3-49　通过 Google 网管工具发出蜘蛛抓取页面

图 3-50　蜘蛛抓取到的 HTML 代码

显然这对站长确认转向设置和检查服务器是否正确返回内容很有帮助。

另外，这个工具也可以用来检查页面是否被黑。有的时候黑客放入的代码会检查浏览器类型，如果是用户使用的普通浏览器访问则返回正常内容，如果是搜索引擎蜘蛛访问，才返回黑客加上去的垃圾内容和垃圾链接。所以站长自己访问网站看不出异样，Google 蜘蛛抓到的却不是站长自己看到的内容。这个工具可以帮助站长检查页面是否存在安全漏洞。

3.12.10　网站性能

网站速度现在越来越被重视，不仅可能对排名产生影响，对用户体验也有很大影响。网管工具网站性能部分显示出网页平均载入时间，如图 3-51 所示。

这个网页打开时间并不是 Google 蜘蛛抓取文件所要花的时间，而是 Google 工具条

记录的普通用户打开页面要花的时间。所以网站服务器地理位置并不影响网站性能部分列出的数据。Google 蜘蛛都是从美国发出的，但网站服务器放在中国不意味着比放在美国载入时间要长，因为 Google 记录的是普通用户访问网站的快慢。

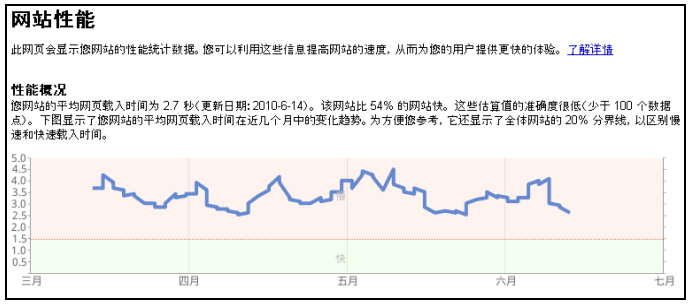


图 3-51　Google 网管工具显示的网站性能

第 4 章
网站结构优化

网站内的优化大致可以分为两部分：一是网站结构调整，二是页面上针对关键词的相关性优化。本章讨论网站结构优化。

网站结构是 SEO 的基础。SEO 人员对页面优化讨论得比较多，如页面上关键词怎样分布，标题标签怎样撰写等，对网站结构的讨论则比较少。其实网站结构的优化比页面优化更重要，掌握起来也更困难。

从 SEO 角度看，优化网站结构要达到以下几个目的。

1．用户体验

用户访问一个网站必须能够不假思索地自如点击链接，找到自己想要的信息。这有赖于良好的导航系统，适时出现的内部链接，准确的锚文字。从根本上说，用户体验好的网站也是搜索引擎喜欢的网站，用户在网站上的行为方式也很可能被计入排名算法中。

2．收录

网站页面的收录在很大程度上依靠良好的网站结构。理论上，清晰的网站结构很容易说清楚，只要策划好分类或频道，然后在分类下加入产品页面，整个网站自然形成树形结构。在实际操作中，大中型网站往往会形成一个异常复杂的链接结构，怎样使搜索引擎蜘蛛能顺利爬行到所有内部页面是一个很大的挑战。

3．权重分配

除了外部链接能给内部页面带来权重外，网站本身的结构及链接关系是内部页面权重分配的重要因素。哪些页面具备比较高的排名能力，取决于页面得到的权重。SEO 人员必须有意识地规划好网站所有页面的重要程度，通过链接结构把权重倾向重要页面。

4．锚文字

锚文字是排名算法很重要的一部分。网站内部链接锚文字是站长自己能控制的，所以是最主要的增强关键词相关性的方法之一。在这方面维基百科是 SEO 人员应该学习的典范，其内部链接及锚文字使用达到了非常高的水平。

本章经常会提到分类、频道、栏目、产品、文章等页面名称。其中分类、频道和栏目页面从网站结构角度看是一回事儿，只不过电子商务网站常称其为分类，门户网站喜欢称其为频道，信息类网站又可能称其为栏目。为了叙述简洁，后面大部分情况下不再

表达为“分类、频道或栏目页面”，只写“分类”页面，但读者应该清楚，优化“分类”页面的技术同样适用于频道和栏目页面，它们在网站结构意义上是完全一样的。产品、文章页面也是如此。

4.1 搜索引擎友好的网站设计

如果我们从搜索引擎蜘蛛的角度去看待一个网站，在抓取、索引和排名时会遇到哪些问题呢？解决了这些问题的网站设计就是搜索引擎友好（search engine friendly）的。

1．搜索引擎蜘蛛能不能找到网页

要让搜索引擎发现网站首页，就必须要有外部链接连到首页，找到首页后，蜘蛛沿着内部链接找到更深的内容页，所以要求网站要有良好的结构，符合逻辑，并且所有页面可以通过可爬行的普通 HTML 链接达到。JavaScript 链接、Flash 中的链接等搜索引擎蜘蛛一般不能跟踪爬行，就会造成收录问题。

网站所有页面离首页点击距离不能太远，最好在四五次点击之内。要被搜索引擎收录，页面需要有最基本的权重，良好的网站链接结构可以适当传递权重，使尽量多的页面达到收录门槛。

2．找到网页后能不能抓取页面内容

被发现的URL必须是可以被抓取的。数据库动态生成、带有过多参数的URL、Session ID、整个页面是 Flash、框架结构（frame）、可疑的转向、大量复制内容等都可能使搜索引擎敬而远之。

某些文件站长可能不希望被收录，除了不链接到这些文件，更保险的方法是使用 robots 文件或 meta robots 标签禁止收录。

3．抓取页面后怎样提炼有用信息

关键词在页面重要位置的合理分布、重要标签的撰写、HTML 代码精简、起码的兼容性等，可以帮助搜索引擎理解页面内容，提取有用信息。这部分在第 5 章“页面优化”部分再深入讨论。

只有搜索引擎能顺利找到所有页面，抓取这些页面并提取出其中真正有相关性的内容，网站才可以被视为是搜索引擎友好的。

网上搜索引擎不友好的网站比比皆是，尤其是一些中小企业网站。不过这样的网站不太好举例，没得到网站主人的许可就当做负面案例讨论非常不恰当。在写本节时刚好在 zaccode.com 网站看到一个会员询问搜索引擎不收录他们的内页是怎么回事。笔者点击过去看了一下，不由得惊呼，这个网站简直是集搜索引擎不友好之大成，是个相当不

错的反面教材。在得到网站所有人正式许可后，在这里作为例子简单介绍。

网站域名是 llyez.com，南昌良良母婴用品有限公司，一个母婴保健及家居用品企业网站。应该说网站视觉设计还是不错的，然而从 SEO 角度看，简直就是个“杯具”。首页抓图如图 4-1 所示。

图 4-1　良良母婴用品网站首页

除了页脚的联系地址、版权声明两行字外，页面其他部分就是一个大的 Flash，源代码如下：

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<title>
南昌良良母婴用品有限公司官网--主营母婴用品 婴幼儿日用品 婴幼儿床上用品 婴儿枕头 尿垫
围嘴 睡袋 凉席
</title>
<link type="text/css" rel="stylesheet" href="../styles/common1.css" />

<meta name="Keywords" content="婴儿枕头，婴幼儿日用品，儿童枕头，保健枕头，婴儿
凉席，围嘴，尿垫，睡袋，宝宝睡眠" />
<meta name="Description" content="南昌良良母婴用品有限公司生产的苎麻系列婴幼儿
用品荣获国家多项专利，被誉为中国婴童保健用品专家。主营婴幼儿日用品、婴幼儿床上用品、
婴儿枕头，儿童枕头，成人枕头，保健枕头等枕头软家居用品，还有婴儿围嘴、凉席、尿垫、睡
袋等婴幼儿日用品。
" />
<!--由中企动力科技集团股份有限公司南昌分公司技术部设计制作<br>如果您有任何意见或建
议请电邮 dm-nanchang@ce.net.cn -->
</head>
<body>

<div class="index-a">
  <object classid="clsid:D27CDB6E-AE6D-11cf-96B8-444553540000"
codebase="http://download.macromedia.com/pub/shockwave/cabs/flash/swfl
ash.cab#version=7,0,19,0" width="980" height="600">
    <param name="movie" value="../images/intro.swf" />
    <param name="quality" value="high" />
        <param name="wmode" value="transparent" />
    <embed src="../images/intro.swf" quality="high"
pluginspage="http: //www.macromedia.com/go/getflashplayer"
type="application/x-shockwave-flash" width="980" height="600"></embed>
```

```
  </object>
<div>
    <div align="center" style="padding-top:25px;">联系地址：江西省南昌市八一大
道 461 号（省医学院内）医科所附三楼
      <br />
  版权所有：南昌良良母婴用品有限公司  <a href="http://nanchang.ce.net.cn"
target="_blank">中企动力提供技术支持 </a>| <a
href="http://www.miibeian.gov. cn/" target="_blank">赣 ICP 备 05010033 号
</a><script src="
http://s9.cnzz. com/stat. php?id=2047204&web_id=2047204&show=pic1"
language="JavaScript"></script></div>
  </div>
</div>
</body>
</html>
```

搜索引擎蜘蛛访问首页后，完全没有任何一个链接能通向内页，倒是有指向网站设计服务商及备案网站的链接。搜索引擎既不能抓取和索引 Flash 中的文字内容，也不能跟踪链接爬到内页。如果不给内部页面直接造些外链的话，整个网站能被收录的基本上只有首页一个页面，其内容还不能被索引。

单击左侧婴儿保健用品链接（Flash 中的链接）进入内页，如图 4-2 所示。

图 4-2　婴儿保健用品页面

看起来设计不错。可惜的是，除了左下角的新闻中心部分，页面上的其他文字、图片还是一个大 Flash，包括顶部的导航条。所以就算站长给这个婴儿用品页面造一些外链，收录又到此为止了。企业概况、商品中心、在线商城等重要部分的页面还是没有任何爬行通路。

左下角的新闻中心设计也有些奇怪，鼠标放上去时才显示链接，而且只显示了一半，如图 4-3 所示。

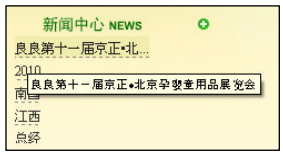


图 4-3　新闻中心链接

点击第一篇新闻中心文章后来到如图 4-4 所示的这个页面。

图 4-4　新闻页面及 URL

顶部导航依然是 Flash，从新闻中心爬到产品页面也不可能了。另一个可怕的地方是 URL，读者可以在抓图地址栏中看到一部分。下面是 HTML 代码的相应部分，读者可以看到完整的网址。

```
<a
href="/InfoContent/id=b49e4b44-5f38-413d-a0c2-bfc2edb7af2e&comp_stats=
comp-FrontInfo_listByAsyncWithOutAjax-123.html" title="良良第十一届京正•
北京孕婴童用品展览会" target="_blank">          良良第十一届京正•北...</a>
```

这种 URL 就算有外部链接，想被收录也难。

从婴儿用品首页单击商品中心来到如图 4-5 所示页面。

图 4-5　商品中心页面及 URL

同样，导航还是 Flash。值得欣慰的是左侧产品链接是普通 HTML 链接，不过单击任何一个产品都可以看到长长的 URL，如图 4-6 所示。

图 4-6　产品页面及 URL

带有这么多参数的 URL，被收录的可能性很低，除非是个权重非常高的网站。这些链接相应的 HTML 代码如下：

```
<div class="abouts_sidebar1">
<h4><a
href="/ProductExhibitlist/&categoryid=bfb98cc0-9890-4e54-b7a7 -26bf85d0
6280&comp_stats=comp-FrontProductCategory_showTree-110.html">芒麻保健枕
</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=56f6b971-9021-4044-bccf -fbece16b
155a&comp_stats=comp-FrontProductCategory_showTree-110.html">芒麻保健袜
</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=f2657c17-ac58-49c6-84cc -3fc14528
0cca&comp_stats=comp-FrontProductCategory_showTree-110.html">芒麻保健凉
席</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=e000a25d-b6c2-40b2-b474 -952ac440
c8b8&comp_stats=comp-FrontProductCategory_showTree-110.html">纯蚕丝被&睡
袋 </a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=c400edfc-2f1a-49de-8b6f -7b01813f
2842&comp_stats=comp-FrontProductCategory_showTree-110.html">芒麻抗菌床
单</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=46507f40-9384-438f-87a0 -f995be9e
7ea6&comp_stats=comp-FrontProductCategory_showTree-110.html">礼包&帽子
</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=9bed3f9c-7a41-439c-be12 -018c4f15
97c6&comp_stats=comp-FrontProductCategory_showTree-110.html">芒麻浴孕用
巾</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=34ef2bd6-ebb6-46d4-907e -9d09225a
ccc9&comp_stats=comp-FrontProductCategory_showTree-110.html">芒麻围嘴食
饭衣</a></h4>
```

```
<h4><a
href="/ProductExhibitlist/&categoryid=d1d7132d-174a-4079-9e0b -dd68f0e7
1b95&comp_stats=comp-FrontProductCategory_showTree-110.html">苎麻抗菌尿
垫床垫</a></h4>
<h4><a
href="/ProductExhibitlist/&categoryid=190a4b2d-02f8-4e20-92d0 -34497acd
333f&comp_stats=comp-FrontProductCategory_showTree-110.html">甲壳素系列
</a></h4>
</div>
```

对比首页、商品中心首页及产品页面标题，大家可以看到这个网站上除了新闻中心的几个页面之外，其他所有页面包括企业概况、商品中心、在线商城、在线调查等页面标题全都一样。

读者看到本书再去访问时，这个网站也许已经不是这个样子了。这是一个典型的搜索引擎不友好的企业网站，只考虑到视觉设计，完全没顾及到怎样被搜索引擎收录，怎样获得搜索流量。读者看完第 4 章和第 5 章，再回头看一遍这个例子，会更了解这个网站为什么搜索引擎不友好，应该怎样改进。

4.2 避免蜘蛛陷阱

有一些网站设计技术对搜索引擎来说很不友好，不利于蜘蛛爬行和抓取，这些技术被称为蜘蛛陷阱。常见的应该全力避免的蜘蛛陷阱包括如下这些。

4.2.1 Flash

在网页的某一小部分使用 Flash 增强视觉效果是很正常的，比如用 Flash 做成的广告、图标等。这种小 Flash 和图片是一样的，只是 HTML 代码中的很小一部分，页面上还有其他以文字为主的内容，所以对搜索引擎抓取和收录没有影响。

但是有的网站整个首页就是一个大的 Flash 文件，这就构成了蜘蛛陷阱。搜索引擎抓取的 HTML 代码只有一个连向 Flash 文件的链接，没有其他文字内容。读者可以参考前面例子的源代码。搜索引擎是无法读取 Flash 文件中的文字内容和链接的。这种网站整个就是一个 Flash 的网站，可能视觉效果做得精彩异常，可惜搜索引擎看不到，不能索引出任何文字信息，也就无从判断相关性。

有的网站喜欢在首页放一个 Flash 动画片头（Flash Intro），用户访问网站看完片头后被转向到真正的 HTML 版本的文字网站首页。搜索引擎不能读取 Flash，一般也没办法从 Flash Intro 跟踪到 HTML 版本页面。

近几年搜索引擎不停尝试读取 Flash 文件，尤其是里面的文字内容和链接，也取得了一定进展。站长可以使用 SWFObject 和 sIFR 等技术使 Flash 文件更易于被搜索引擎读取。不过我的观点是，与其费劲去研究怎么让搜索引擎读取 Flash 文件，不如做成普通的 HTML 版本的网站。如果你的网站除了 Flash 外不足以展现公司实力，那么你的公司应该是世界 500 强级别的，SEO 就不重要了。

整个网站就是一个大 Flash 文件，从 SEO 角度来说是万万不可的。如果 Flash 效果是必需的，至少也需要在首页加上一个通往 HTML 版本的链接。这个链接应该是在 Flash 文件之外的 HTML 代码中，搜索引擎跟踪这个链接可以抓取后面的 HTML 版本页面。

4.2.2 Session ID

有些网站使用 Session ID（会话 ID）跟踪用户访问，每个用户访问网站时都会生成独特唯一的 Session ID，加在 URL 中。搜索引擎蜘蛛的每一次访问也会被当成一个新的用户，URL 中会加上一个不同的 Session ID，这样搜索引擎蜘蛛每次来访问时所得到的同一个页面的 URL 将不一样，后面带着一个不一样的 Session ID。这也是最常见的蜘蛛陷阱之一。

搜索引擎遇到这种长长的 Session ID，会尝试判断字符串是 Session ID 还是正常参数，成功判断出 Session ID 就可以去掉 Session ID，收录正常 URL。但也经常判断不出来，要么不愿意收录，要么收录多个带有不同 Session ID 的 URL，内容却完全一样，形成大量复制内容，这两种情况对网站优化都不利。

通常建议跟踪用户访问应该使用 cookies 而不要生成 Session ID。或者程序判断访问者是搜索引擎蜘蛛还是普通用户，如果是搜索引擎蜘蛛，则不生成 Session ID。跟踪搜索引擎蜘蛛访问是没什么意义的，蜘蛛既不会填表，也不会把商品放入购物车。

4.2.3 各种跳转

除了后面会介绍的 301 转向以外，搜索引擎对其他形式的跳转都比较敏感，如 302 跳转、JavaScript 跳转、Flash 跳转、Meta Refresh 跳转。

有些网站用户访问首页时会被自动转向到某个目录下的页面。如果是按用户地理位置转向至最适合的目录，那倒情有可原。但大部分这种首页转向看不出任何理由和目的，这样的转向能避免则避免。

如果必须转向，301 转向是搜索引擎推荐的、用于网址更改的转向，可以把页面权重从旧网址转移到新网址。其他转向方式都对蜘蛛爬行不利，原因在于后面作弊一章所要介绍的，黑帽 SEO 经常使用转向欺骗搜索引擎和用户。所以除非万不得已，尽量不要使用 301 转向以外的跳转。

4.2.4 框架结构

如果作为站长的你不知道什么是框架结构（Frame），那么恭喜你，你已经避免了这个蜘蛛陷阱，根本没必要知道什么是框架结构。如果你在网站设计中还在使用框架结构，我的建议是立即取消。

使用框架结构设计页面，在网站诞生初期曾经挺流行，因为对网站的页面更新维护有一定方便性。不过现在的网站已经很少使用框架了，不利于搜索引擎抓取是框架越来越不流行的重要原因之一。对搜索引擎来说，访问一个使用框架的网址所抓取的 HTML 只包含调用其他 HTML 文件的代码，并不包含任何文字信息，搜索引擎根本无法判断这个网址的内容是什么。虽然蜘蛛可以跟踪框架中所调用的其他 HTML 文件，但是这些文件经常是不完整的页面，比如没有导航只是正文。搜索引擎也无法判断框架中的页面内容应该属于主框架，还是属于框架调用的文件。

总之，如果你的网站还在使用框架，或者你的老板要使用框架结构，唯一要记住的是，忘记使用框架这回事。和 Flash 一样，别浪费时间研究怎么让搜索引擎收录框架结构网站。

4.2.5 动态 URL

动态 URL 指的是数据库驱动的网站所生成的、带有问号、等号及参数的网址。一般来说动态 URL 不利于搜索引擎蜘蛛爬行，应该尽量避免。下面对动态 URL 及其静态化还有更详细的讨论。

4.2.6 JavaScript 链接

由于 JavaScript 可以创造出很多吸引人的视觉效果，有些网站喜欢使用 JavaScript 脚本生成导航系统。这也是比较严重的蛛蛛陷阱之一。虽然搜索引擎都在尝试解析 JS 脚本，不过我们不能寄希望于搜索引擎自己去克服困难，而要让搜索引擎跟踪爬行链接的工作尽量简单容易。

据我观察，虽然有的搜索引擎在技术上可以获得 JavaScript 脚本中包含的链接，甚至可以执行脚本并跟踪链接，但对一些权重比较低的网站，搜索引擎觉得没有必要，不会费那个劲。所以网站上的链接必须使用最简单标准的 HTML 链接，尤其是导航系统。用 CSS 做导航系统一样可以实现很多视觉效果。

JavaScript 链接在 SEO 中也有特殊用途，那就是站长不希望被收录的页面（比如没有排名意义的页面，重复内容页面等），不希望传递权重的链接，可以使用 JavaScript 脚本阻挡搜索引擎爬行。

4.2.7 要求登录

有些网站内容放在需要用户登录之后才能看到的会员区域，这部分内容搜索引擎无法看到。蜘蛛不能填写用户名、密码，也不会注册。

4.2.8 强制使用 Cookies

有些网站为了实现某种功能，如记住用户登录信息、跟踪用户访问路径等，强制用户使用 Cookies，用户浏览器如果没有启用 Cookies，页面显示不正常。搜索引擎蜘蛛就相当于一个禁用了 Cookies 的浏览器，强制使用 Cookies 只能造成搜索引擎蜘蛛无法正常访问。

4.3 物理及链接结构

网站结构有两方面的意思：一是物理结构，二是链接结构。

4.3.1 物理结构

网站物理结构指的是网站真实的目录及文件所在的位置所决定的结构。

一般来说比较合理的物理结构有两种。一种是扁平式的，也就是所有网页文件都存在网站根目录下：

- http://www.domain.com/index.html
- http://www.domain.com/catA.html
- http://www.domain.com/catB.html
- http://www.domain.com/pageA.html
- http://www.domain.com/pageB.html
- http://www.domain.com/pageC.html

……

这比较适合小型网站。如果有太多文件放在根目录下，制作和维护起来都比较麻烦，容易搞乱。大中型网站把文件都放在根目录下基本上就无法管理了。

扁平物理结构的一个优势是，很多人认为根目录下的文件比深层目录中的文件天生权重高一点。比如，如果其他条件相同，URL：

http://www.domain.com/pageA.html

比 URL：

http://www.domain.com/catA/pageA.html

排名能力高一点。

第二种是树形结构，或称为金字塔形结构。根目录之下以目录形式分成多个产品分

类（或称为频道、类别、目录、栏目等），然后在每一个分类下再放上属于这个分类的具体产品（或称为文章、帖子等）页面。

比如分类分为：

- http://www.domain.com/catA/
- http://www.domain.com/catB/
- http://www.domain.com/catC/

……

分类下再放入具体的产品页面：

- http://www.domain.com/catA/product-a.html
- http://www.domain.com/catA/product-b .html
- http://www.domain.com/catA/product-c .html
- http://www.domain.com/catB/product-d .html
- http://www.domain.com/catB/product-e .html
- http://www.domain.com/catC/product-f .html

……

树形结构逻辑清晰，页面之间的隶属关系一目了然。

需要说明的是，这里说物理结构是“真实的目录及文件所在的位置所决定的结构”，只是为了方便说明和理解。数据库驱动、程序生成的网站并不存在真实的目录和文件，URL 中的目录和文件都是程序实时生成的，但就网站结构来说，与真实存在的目录和文件没有什么区别。

4.3.2 链接结构

网站结构的第二个意义指的是链接结构，或称为逻辑结构，也就是由网站内部链接形成的链接的网络图。

比较合理的链接结构通常是树形结构，如图 4-7 所示。

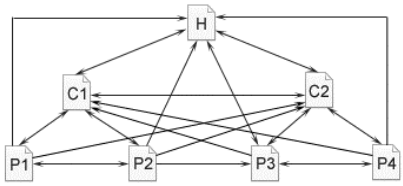


图 4-7　树形网站链接结构示意图

H 为网站首页，C1 和 C2 是分类首页，P1、P2 是分类 C1 下产品页面，P3、P4 是分类 C2 下产品。其中的链接关系如下：

- 首页链接向所有分类首页。
- 首页一般不直接链接向产品页，除了几个需要特殊推广的产品，如 P3。
- 所有分类首页连向其他分类首页，一般以网站导航形式体现。
- 分类首页都链接回网站首页。
- 分类首页链接向本分类下的产品页。
- 分类首页一般不链接向其他分类的产品页。
- 产品页都链接向网站首页，一般以网站导航形式体现。
- 产品页链接向所有分类首页，一般以网站导航形式体现。
- 产品页可以链接向同一个分类的其他产品页。
- 产品页一般不链接向其他分类的产品页。
- 在某些情况下，产品页可以用适当的关键词链接向其他分类的产品页，如 P2 链接向 P3。

从图示和说明可以看到，这些链接会很自然地形成一个树形的网络图。这种链接网络可以与物理结构重合，也可以不一样，比如扁平式的物理结构网站完全可以通过链接形成链接上的树形结构。

当然实际网站的链接结构要复杂得多，一级分类不可能只有两个，下面还可能有二级、三级分类，末级分类可能有很多个翻页，还可能有各种排序页面等。但网站链接基本形式大体相同。

对搜索引擎来说更重要的是链接结构，而不是物理结构。不少人有误解，认为物理结构比较深的页面不容易被搜索引擎收录，比如：

http://www.domain.com/cat1/cat1-1/cat1-1-1/pageA.html

像这样物理目录结构比较深的 URL，是不是就不容易被收录呢？那不一定。如果这个页面在网站首页上有一个链接，对搜索引擎来说它就是一个仅次于首页的链接结构意义上的二级页面。收录容易与否在于页面处于链接结构的什么位置，离首页有几次点击距离，而不是它的目录层次。

树形链接结构使权重在网站各页面上均匀分布，深层内页可以从首页 4～5 次点击内达到。但当网站规模比较大时，还是会出现各种问题，第 4.14 节将着重讨论对树形链接结构的补充和修改。

4.4 清晰导航

清晰的导航系统是网站设计的重要目标，对网站信息架构、用户体验影响重大。SEO 也越来越成为导航设计时需要考虑的因素之一。

站在用户角度，网站导航系统需要解决两个问题。

（1）**我现在在哪里**？用户可能从任何一个内页进入网站，有的时候从首页进入，点击多个链接后，用户也已经忘了是怎么来到当前页面的。导航系统这时就要清楚地告诉用户现在处在网站总体结构的哪一部分。

页面设计风格的统一、面包屑导航的使用、主导航系统当前所在分类高亮都有助于用户判断自己现在在哪里。

（2）**下一步要去哪里**？有的时候用户知道自己想做什么，页面的导航设计要告诉用户点击哪里才能完成他的目标。有的时候用户也不知道自己该干什么，网站导航就要给用户一个最好的建议，引导用户流向网站目标完成页面。

合理的导航及分类名称、正文中的相关链接、引导用户把产品放入购物车的按钮、相关产品推荐、网站地图、站内搜索框等都有助于帮助用户点击到下一步。

站在 SEO 的角度，网站导航系统应该注意以下几点。

1．文字导航

尽量使用最普通的 HTML 文字导航，不要使用图片作为导航链接，更不要使用 JavaScript 生成导航系统，也不要用 Flash 做导航。CSS 也可以设计出很好的视觉效果，如背景、文字颜色变化，下拉菜单等。

最普通的文字链接对搜索引擎来说是阻力最小的爬行抓取通道。导航系统链接是整个网站收录最重要的内部链接，千万不要在导航上给搜索引擎设置任何障碍。

2．点击距离及扁平化

良好导航的目标之一是使所有页面与首页点击距离越近越好。权重普通的网站，内页离首页不要超过四五次点击。要做到这一点，通常应该在链接结构上使网站尽量扁平化。

网站导航系统的安排对减少链接层次至关重要。主导航中出现的页面将处于仅低于首页的层次，所以主导航中页面越多，网站越扁平。但用户体验和页面链接总数都不允许主导航中有太多链接。SEO 人员需要在网站规模、用户体验中做好平衡。

3．锚文字包含关键词

导航系统中的链接通常是分类页面获得内部链接的最主要来源，数量巨大，其锚文字对目标页面相关性有相当大的影响，因此分类名称应尽量使用目标关键词。当然首先要顾及用户体验，导航中不能堆积关键词，分类名称以 2～4 个字为宜。

4．面包屑导航

面包屑导航对用户和搜索引擎来说，是判断页面在网站整个结构中的位置的最好方

法。正确使用面包屑导航的网站通常都是架构比较清晰的网站，强烈建议使用，尤其是大中型网站。

5．避免页脚堆积

随着 SEO 被更多站长认识，近几年有一种在页脚堆积富含关键词的分类页面链接的倾向。两三年前这种做法还有不错的效果，但是近来搜索引擎比较反感这种做法，常常造成某种形式的惩罚。

4.5 子域名和目录

搜索引擎通常会把子域名（或称为二级域名）当做一个基本独立的站点看待，也就是说

http://www.domain.com

和

http://news.domain.com

是两个互相独立的网站。

而目录

http://www.domain.com/news/

就纯粹是 http://www.domain.com 的一部分。

顺便说一句，www.domain.com 其实是 domain.com 的一个子域名，只是约定俗成，网站一般以 www.domain.com 为主 URL。

如果抛开其他因素，只看这两个 URL：

- http://news.domain.com
- http://www.domain.com/news/

子域名 http://news.domain.com 的权重稍微高一点，因为搜索引擎会把这个 URL 当做网站的首页。另外，很多人观察到，主域名很多时候会传递一小部分信任度（是信任度，不是 PR）给子域名。

所以单就 URL 来看，子域名比目录天生的权重和排名能力稍微高一点。但从 SEO 角度看，我建议尽可能使用目录，而不是子域名。

子域名和主域名是两个完全不同的网站，要推广的也是两个网站，所有的优化工作都要多做一遍，尤其是外部链接建设，网站 PR 值、权重都会被这两个独立的网站分散。

主域名经过外链建设获得高权重，不意味着子域名就获得了高权重。

子域名的使用会使网站变多，同时使每个子域名网站变小。目录会使一个网站越做越大。网站越大，包含的内容自然就越多，对用户的帮助越大，它所累积的信任度就更高。

当然，这只是对一般网站而言的。在某些情况下，子域名是更适当的。比如：

- 网站内容足够多，每一个子域名下涵盖的内容都足以成为一个独立的网站而毫不逊色。有很多这样的网站，诸如各类门户。像新浪、搜狐这种级别的网站，任何一个频道的内容都比绝大多数网站多得多。
- 跨国公司不同国家的分部或分公司，采取子域名有利于建立自己品牌。而且各个分公司的网站内容很可能是由不同国家的团队自行维护的，与独立网站没什么区别。
- 公司有不同的产品线，而且相互之间联系不大，或者完全是以不同的品牌出现的，这时每一个品牌或产品线可以用子域名甚至独立域名。如几乎所有汽车公司每一个品牌甚至每一个车型都有自己独立的品牌和网站。手机公司也如此。门户类网站推出不同服务时，如博客、微博、论坛、邮件，也以子域名为宜。
- 分类信息网站，不同城市使用不同的子域名。这类分类信息网站通常都有海量数据，某一个省份或城市的内容都能独立成站。而且不同城市、省份很可能有不同的发展目标和安排，管理也是由不同的团队操作的。
- 平台类网站，如 B2B 和淘宝类的商城。每个用户都有自己相对独立的展示平台，使用自己的子域名对品牌建设、推广有好处。

除了网站规模和类型，子域名和目录的选择有时还得看其他情况。比如，目标搜索引擎如果是百度，使用子域名或独立域名更有利，因为百度给网站首页权重更高。相比之下，Google 给予网站首页、目录及内页大致相同的机会。如果是做英文网站，采用目录结构，不会出现明显的不利。

4.6 禁止收录机制

有的时候，站长并不希望某些页面被抓取和收录，如付费内容、还在测试阶段的页面、复制内容页面等。网站上不出现链接，或者使用 JavaScript、Flash 链接，使用 nofollow 等方法都不能保证页面一定不被收录。站长自己虽然没有链接到不想被收录的页面，其他网站可能由于某种原因出现导入链接，导致页面被收录。

要确保页面不被收录，需要使用 robots 文件或 Meta Robots 标签。

4.6.1 robots 文件

搜索引擎蜘蛛访问网站时，会先查看网站根目录下有没有一个命名为 robots.txt 的纯文本文件，robots.txt 用于指令搜索引擎禁止抓取网站某些内容或指定允许抓取某些内容。

如百度的 robots 文件位于：

http://www.baidu.com/robots.txt

只有在需要禁止抓取某些内容时，写 robots.txt 才有意义。robots 文件不存在或者是空文件都意味着允许搜索引擎抓取所有内容。有的服务器设置有问题，robots 文件不存在时会返回 200 状态码及一些错误信息，而不是 404 状态码，这有可能使搜索引擎错误解读 robots 文件信息，所以建议就算允许抓取所有内容，也要建一个空的 robots.txt 文件，放在根目录下。

robots 文件由记录组成，记录之间以空行分开。记录格式为：

<域>:<可选空格><域值><可选空格>

最简单的 robots 文件：

User-agent: *

Disallow: /

上面这个 robots 文件禁止所有搜索引擎抓取任何内容。

User-agent:指定下面的规则适用于哪个蜘蛛。通配符* 代表所有搜索引擎。只适用于百度蜘蛛则用：

User-agent: Baiduspider

只适用于 Google 蜘蛛则用：

User-Agent: Googlebot

Disallow: 告诉蜘蛛不要抓取某些文件或目录。例如下面的代码将阻止所有蜘蛛抓取/cgi-bin/和/tmp/两个目录下的内容及文件/aa/index.html：

```
User-agent: *
Disallow: /cgi-bin/
Disallow: /tmp/
Disallow: /aa/index.html
```

Disallow：禁止的目录或文件必须分开写，每个一行，不能写成：

Disallow: /cgi-bin/ /tmp/ /aa/index.html

下面的指令相当于允许所有搜索引擎抓取任何内容：

```
User-agent: *
Disallow:
```

下面的代码禁止除百度外的所有搜索引擎抓取任何内容：

```
User-agent: Baiduspider
Disallow:

User-agent: *
Disallow: /
```

Allow: 告诉蜘蛛应该抓取某些文件。由于不指定就是允许抓取，Allow:单独写没有意义，Allow 和 Disallow 配合使用，可以告诉蜘蛛某个目录下大部分不允许抓取，只允许抓取一部分。例如，下面的代码将使蜘蛛不抓取/ab/目录下其他目录和文件，但允许抓取其中/cd/目录下的内容：

```
User-agent: *
Disallow: /ab/
Allow: /ab/cd/
```

$通配符：匹配 URL 结尾的字符。例如，下面的代码将允许蜘蛛抓取以.htm 为后缀的 URL：

```
User-agent: *
Allow: .htm$
```

下面的代码将禁止百度抓取所有.jpg 文件：

```
User-agent: Baiduspider
Disallow: .jpg$
```

*通配符：告诉蜘蛛匹配任意一段字符。例如，下面一段代码将禁止蜘蛛抓取所有 htm 文件：

```
User-agent: *
Disallow: /*.htm
```

Sitemaps 位置：告诉蜘蛛 XML 网站地图在哪里，格式为：

Sitemap: <网站地图位置>

主流搜索引擎都遵守 robots 文件指令，robots.txt 禁止抓取的文件搜索引擎将不访问，不抓取。但要注意的是，被 robots 文件禁止抓取的 URL 还是可能出现在搜索结果中。只要有导入链接指向这个 URL，搜索引擎就知道这个 URL 的存在，虽然不会抓取页面内容，但是可能以下面几种形式显示在搜索结果中：

- 只显示 URL，没有标题、描述。Google 常这样处理。
- 显示开放目录或雅虎等重要目录收录的标题和描述。
- 导入链接的锚文字显示为标题和描述。百度常这样处理。
- 要想使 URL 完全不出现在搜索结果中，需要使用页面上的 meta robots 标签。

4.6.2 meta robots 标签

meta robots 标签是页面 head 部分 meta 标签的一种，用于指令搜索引擎禁止索引本页内容。

最简单的 meta robots 标签格式为：

<meta name="robots" content="noindex,nofollow">

上面标签的意义是禁止所有搜索引擎索引本页面，禁止跟踪本页面上的链接。

Google、必应、雅虎都支持的 meta robots 标签如下。

- NOINDEX：告诉蜘蛛不要索引本页面。
- NOFOLLOW：告诉蜘蛛不要跟踪本页面上的链接。
- NOSNIPPET：告诉搜索引擎不要在搜索结果中显示说明文字。
- NOARCHIVE：告诉搜索引擎不要显示快照。
- NOODP：告诉搜索引擎不要使用开放目录中的标题和说明。

百度的官方说法是目前只支持 NOFOLLOW 和 NOARCHIVE。

meta robots 标签内容可以写在一起，以逗号间隔，中间可以有空格，也可以没有。多个 meta robots 内容也可以写成不同标签。

<META NAME="ROBOTS" CONTENT="NOINDEX">

<META NAME="ROBOTS" CONTENT="NOFOLLOW">

与下面这个是一样的：

<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">

meta robots 标签不区分大小写。

只有禁止索引时，使用 meta robots 才有意义。以下这个标签：

<META NAME="ROBOTS" CONTENT="INDEX, FOLLOW">

是没有意义的，等于没写。普通需要被收录、索引，链接需要被跟踪的页面，不用写 meta robots 标签。

这个标签有时会用到：

<meta name="robots" content="noindex">

效果是禁止索引本页面，但允许蜘蛛跟踪页面上的链接，也可以传递权重。

使用了 noindex meta robots 标签的页面会被抓取，但不会被索引，页面 URL 不会出现在搜索结果中，这一点与 robots 文件不同。

4.7 nofollow 的使用

nofollow 是 2005 年由 Google 领头新创的一个标签（严格说是属性），目前主流搜索引擎包括百度、雅虎、必应等都已支持。

nofollow 代码形式为：

这里是锚文字

链接的 nofollow 属性只适用于本链接。上一节介绍的 Meta Robots 标签中的 nofollow 指的是页面上的所有链接。

nofollow 最初目的是减少垃圾链接对搜索引擎排名的影响，标签意义是告诉搜索引擎这个链接不是经过站长自己编辑的，所以这个链接不是一个信任投票。搜索引擎看到这个标签就不会跟踪爬行链接，也不传递链接权重和锚文字。

nofollow 标签通常用在博客评论、论坛帖子、社会化网站、留言板等地方，因为在这些地方任何用户可以自由留下链接，站长一般并不知道链接连向什么网站，也不可能一一查看验证，所以是垃圾链接最多出现的地方。现在主流的博客和论坛软件都自动在评论和帖子的链接中加上了 nofollow。

后来 nofollow 又有了另外一个用——广告链接。网络广告的初衷应该是曝光率和点击流量，而不是 PR 值或试图影响搜索引擎排名。但广告同时也是个链接，会影响权重流动和搜索引擎排名。搜索引擎对试图影响、操纵排名的链接买卖是深恶痛绝的。给广告链接加上 nofollow 可以告诉搜索引擎，这是个广告，请不要传递链接权重。

由于 nofollow 属性能阻止蜘蛛爬行和传递权重，因此很快又被 SEO 用在某些内部链接，以达到控制内部链接权重和 PR 流动及分布的目的。最常见的应用是在联系我们、隐私权政策、用户条款、用户登录等链接上。这些页面往往有整站链接，如果没有使用 nofollow，整站的链接权重会平等地流动到这些网页上。而像隐私权政策这种网页一般来说用户很少关心，想通过搜索排名带来流量的可能性也极低。所以流动到这些搜索价值不高网页的权重浪费了很大一部分。

nofollow 还可以用于其他搜索价值不大因而不需要权重的页面，如电子商务网站产品列表的各种排序方式、显示格式等。

有的网站通过 nofollow 控制权重流动，将权重导向需要收录和排名的页面后，整体

收录、排名和流量有明显提高。但有时候没有明显效果。

也有人认为，用 nofollow 控制链接和权重流动是个迷思，有欺骗搜索引擎之嫌，因为展现给用户和搜索引擎的内容不一样。提高网站内部链接效率依赖于合理的网站架构，而不是使用 nofollow 标签。

Google 资深工程师 Matt Cutts 曾对此评论说，nofollow 使用在内部链接上，确实能够影响 Google 排名，但这只是次要因素。好比你有 100 块钱，使用 nofollow 就好像仔细研究怎么花这 100 块钱，对一些人有帮助。但更有用的是把精力放在怎么再多挣 300 块钱，而不是琢磨怎么花这现有的 100 块钱。

2008 年上半年的某个时间，Google 对 nofollow 链接权重传递算法做了改变。假设页面 PR 和权重有 10 块钱可以传递，页面上有 10 个链接，在 nofollow 标签出现之前，每个链接传递 1 块钱。如果其中 5 个链接加了 nofollow，2008 年的算法改变之前，剩下的没有 nofollow 的 5 个链接，每个链接传递 2 块钱，改变之后的处理方式是，剩下的 5 个非 nofollow 链接还是传递 1 块钱。当然，改变前后，nofollow 链接都不传递 PR。也就是说，2008 年之后，加了 nofollow 相当于浪费了 PR 和权重，页面本来有的 10 块钱，只传递出去 5 块，还有 5 块消失了。

Google 在做了这个变动一年多之后，2009 年 6 月 Matt Cutts 才在博客上宣布。在这一年多时间里，没有人注意到这个算法变动。这说明 nofollow 的具体细节和效果远不是 SEO 人员所能掌握的。更宽泛点看，页面链接怎样分配和传递权重很可能更为复杂，并不是平均分配的。比如，导航及博客评论等地方的链接能传递的权重可能本来就很少。就算没有 nofollow 的存在，博客评论里的链接传递与正文里的链接一样的权重本来就不太合理。

其他搜索引擎怎样处理 nofollow 链接权重分配，目前还是未知的。

4.8 URL 静态化

URL 静态化一直以来都是最基本的 SEO 要求之一，但近一两年 SEO 行业对是否一定要做静态化有了一些观念上的改变。

4.8.1 为什么静态化

现在的网站绝大多数是数据库驱动，页面由程序实时生成，而不是真的在服务器上有一个静态 HTML 文件存在。当用户访问一个网址时，程序根据 URL 中的参数调用数据库数据，实时生成页面内容。因此动态页面相对应的 URL 原始状态也是动态的，包含问号、等号及参数，如下面这种典型论坛的 URL：

http://www.domain.com/viewthread.php?tid=70376&extra=page=1

搜索引擎在发展初期（其实也就是 10 年前到三四年前而已）一般不太愿意爬行和收录动态 URL，主要原因是可能陷入无限循环或收录大量重复内容，造成资源极大浪费。最典型的无限循环就是某些网站上出现的万年历，很多博客都按时间存档，一些宾馆、航班查询网站都经常出现万年历形式，如图 4-8 所示。

图 4-8　万年历使蜘蛛可以无限点击下去

搜索引擎蜘蛛碰到万年历，如果一直跟踪上面的链接，可以不停地点击下一月、下一年，无穷无尽陷入无限循环，每一个日期对应的页面内容也没什么区别。真实用户一眼就能看出这是个日历，但搜索引擎蜘蛛面对的只是一串代码，不一定能判断出其实这是个万年历。

有时就算不存在无限循环，动态 URL 也可能造成大量复制页面。比如 URL ：

http://www.domain.com/product.php?color=red&cat=shoes&id=12345

和 URL：

http://www.domain.com/product.php?cat=shoes&color=red&id=12345

及 URL：

http://www.domain.com/product.php?color=red&id=12345&cat=shoes

很可能是完全一样的内容，都是型号为 12345 的红色鞋子。URL 中参数顺序不同就是不同的网址，但调用参数一样，因此页面内容是一样的。如果 CMS 系统设计不周全，这些 URL 都可能出现在网站上。

更麻烦的是，有时某些参数可以是任意值，服务器都能正常返回页面，虽然内容全是一样或非常相似的。例如，上面 URL 中的参数 12345 改为 6789 或其他什么数字，服务器很可能也返回 200 状态码。

所以以前的搜索引擎对动态 URL 敬而远之，要想网站页面被充分收录，站长需要把动态 URL 转化为静态 URL。

4.8.2 怎样静态化 URL

最常见的方法是使用服务器的 URL 重写模块，在 LAMP（Linux+Apache+MySQL+PHP）服务器上一般使用 mod_rewrite 模块，Windows 服务器也有功能相似的 ISAPI Rewrite 等模块。以 LAMP 服务器为例，要想把 URL：

http://www.domain.com/products.php?id=123

静态化为：

http://www.domain.com/products/123

需要启用服务器 mod_rewrite 模块，然后在.htaccess 文件中写入如下代码：

RewriteRule /products/([0-9]+) /products.php?id=$1

URL 重写代码基于正则表达式，每个网站的动态 URL 结构不同，服务器设置也可能不同，代码也就不同。正则表达式的写法比较复杂，千变万化，通常需要程序员编写。在写 URL 静态化代码时必须非常小心，错了一个字符，多了一个斜杠等微小的不同，都可能造成灾难性的后果。我本人就遇到过这样的案例。

严格来说，这里所说的 URL 静态化应该称为“伪静态化”，也就是说服务器上还是不存在相应的 HTML 文件，用户访问时还是动态生成页面，只不过通过 URL 重写技术使网址看起来像是静态的。也有的 CMS 系统可以实现真正静态化，站长增添产品或文章后，系统会自动生成真实存在的静态的 HTML 文件。对搜索引擎来说，真正的静态与伪静态没有区别。

4.8.3 URL 不需要静态化吗

近两三年搜索引擎对动态 URL 的抓取有了很大进步。一般来说 URL 中有两三个参数，对收录不会造成多大影响。权重高的域名，再多几个问号也不是问题。不过一般来说还是建议将 URL 静态化，既能提高用户体验，又能降低收录难度。

2008 年 9 月，Google 站长博客发表了一篇讨论动态网址和静态网址的帖子，颠覆了这个 SEO 界的传统观念。在这篇帖子里，Google 明确建议不要将动态 URL 静态化。

Google 的帖子有几个要点。

（1）Google 完全有能力抓取动态网址，多少个问号都不是问题。

（2）动态网址更有助于 Google 蜘蛛读懂 URL 含义并进行鉴别，因为网址中的参数有提示性。Google 举了这个例子：

www.example.com/article/bin/answer.foo?language=en&answer=3&sid=98971298178906&query=URL

URL 里的参数本身有助于 Google 理解 URL 及网页内容，比如 language 后面跟的参数是提示语言，answer 后面跟的是文章编号，sid 后面的肯定是 Session ID。其他常用的参数包括：color 后面跟的一般是颜色，size 后面跟的参数是尺寸等。有了这些参数名称的帮助，Google 更容易理解网页。

将网址静态化后，这些参数的意义通常就变得不明显了，比如这个 URL：

www.example.com/shoes/red/7/12/men/index.html

就可能使 Google 不知道哪个是产品序列号，哪个是尺寸。

（3）URL 静态化很容易弄错，那就更得不偿失了。比如，通常动态网址的参数调换顺序所得到的页面是相同的，这两个网址很可能就是同一个页面：

- www.example.com/article/bin/answer.foo?language=en&answer=3
- www.example.com/article/bin/answer.foo?answer=3&language=en

保留动态网址，Google 比较容易明白这两个 URL 是一个页面，因而自动合并权重。经过静态化后，这样两个网址 Google 就不容易判断是不是同一个页面，从而可能引起复制内容：

- www.example.com/shoes/men/7/red/index.html
- www.example.com/shoes/red/7/men/index.html

再一个容易搞错的是 Session ID，也可能被静态化进 URL：

www.example.com/article/bin/answer.foo/en/3/98971298178906/URL

这样网站将产生大量 URL 不同，但其实内容相同的页面。

所以，Google 建议不要静态化 URL。

但是我目前还是建议尽量要静态化 URL。原因如下。

（1）Google 给的建议是从 Google 自己出发，没有考虑其他搜索引擎。Google 抓取任何动态网址没问题，并不意味着百度、雅虎、必应、搜狗、搜搜等都没问题。

（2）Google 所说的静态化的弊端，基本上是基于静态化做得不正确的假设。问题是要做静态化就得做正确，假设会做错是没有什么道理的。有几个人会在静态化网址时把 Session ID 也放进去呢？

（3）用户体验。带有参数的 URL 可能有助于 Google 读懂内容，但是显然非常不利于用户在一瞥之下理解页面大致内容。看看这两个网址哪个更清晰，更容易读懂，更有可能被点击呢：

- www.example.com/product/bin/answer.foo?language=en&productID=3&sid=98971298178906&cat=6198&&query=URL
- www.example.com/product/men/shoes/index.html

显然是第二个。

除了点击率，长的动态网址也不利于记忆，不利于在邮件、社会化网站等地方抄送给别人。

4.9 URL 设计

URL 在搜索结果列表中是显示内容之一。设计网站结构时需要对目录及文件命名系统做事先规划。总的原则是首先从用户体验出发，URL 应该清晰友好、方便记忆，然后才考虑 URL 对排名的影响。具体可以考虑以下几方面。

1．URL 越短越好

这主要是为用户着想。对搜索引擎来说，只要 URL 不超过 1000 个字母，收录起来都没问题。不过真的使用几百个字母的 URL，用户看起来就费事了。曾经有人做过搜索结果点击实验，一个比较短的 URL 出现在一个比较长的 URL 下面时，短 URL 的点击率比长 URL 高 2.5 倍。

另外，短 URL 也利于传播和复制。站长做链接时，通常会直接复制 URL。短 URL 不会有问题，长 URL 复制时都会费劲，也可能复制得不完整，造成 404 错误。

2．避免太多参数

在可能的情况下尽量使用静态 URL。如果技术上不能实现，必须使用动态 URL，也要尽量减少参数。一般建议 2～3 个参数之内。参数太多用户看着眼花缭乱，也可能造成收录问题。

3．目录层次尽量少

这里指的是物理目录结构。

当然目录层次与网站整个分类结构相关。分类层数越多，目录层次也必然增多。在可能的情况下，尤其是静态化 URL 时，尽量使用比较少的目录层次。根据观察，百度尤其喜欢目录层次少的页面。

当然这不是说建议大家把页面全放在根目录下，那样的话，超过几百页的网站就不容易管理了，不仅搜索引擎无法从目录层次了解归属关系，站长自己恐怕也不容易分清哪个页面属于哪个分类。

4．文件及目录名具描述性

尤其对英文网站来说，目录及文件名应该具备一定的描述性，使用户在一瞥之下就能知道这个 URL 内容大致应该是什么。比如 URL http://www.example.com/news/finance/ 就比 http://www.example.com/cat-01/sub-a/要好得多。

5．URL 中包含关键词

关键词出现在 URL 中，也能提高页面相关性，在排名时贡献一点分数。关键词出现得越靠前越好，也就是说出现在域名中最好，其次是出现在目录名中，效果最小的是出现在文件名中。不过切不可为了出现关键词而堆积。

6．字母全部小写

这有以下几方面原因。

（1）全部小写容易人工输入，不会因大小写掺杂而犯错。

（2）有的服务器是区分大小写的，如 Linux 服务器，也就是说 http://www.example.com/index.html 与 http://www.example.com/Index.html。

以上是两个不同的网址。无论站长自己在做链接时还是用户输入时，因为大小写混用出现错误都会造成 404 错误。另外，robots 文件代码也是区分大小写的，一个字母之差就可能使整个目录不能被收录。

本书举 URL 例子时有时用大写字母，只是为了强调，让读者看得更清楚。

7．连词符使用

目录或文件名中单词间一般建议使用短横线（- ）分隔，不要使用下画线或其他更奇怪的字母。搜索引擎把 URL 中的短横线当做空格处理，下画线则被忽略。所以文件名 seo-tools.html 将被正确读取出 seo 与 tools 两个单词，而文件名 seotools.html 就不能被分解出两个单词。

8．目录形式还是文件形式

大部分 CMS 系统都可以把页面 URL 设置为目录或文件形式，比如 http://www.example/ products/red-shoes/或 http://www.example/products/red-shoes.html。

站长可以选择其中之一。这两种格式对排名没有大影响。有人认为目录形式的权重稍微高一点，不过也无法验证，就算权重高一点，也应该是微乎其微的。

目录形式的一个优点是，以后如果网站更换编程语言，URL 可以不必变化，也不用经过特殊处理。文件形式的 URL 可能文件扩展名会变化，除非做 URL 重写。

4.10 网址规范化

网址规范化（URL canonicalization）指的是搜索引擎挑选最合适的 URL 作为真正（规范化的）网址的过程。

4.10.1 为什么出现不规范网址

举例来说，下面这几个 URL 一般指的是同一个文件：

- http://www.domainname.com
- http://domainname.com
- http://www.domainname.com/index.html
- http://domainname.com/index.html

但是从技术上来讲，这几个 URL 都是不同的网址，搜索引擎也确实把它们当做不同的网址。虽然在绝大部分情况下，这些网址所返回的是相同的文件，也就是网站首页，但是从技术上说，主机完全可以对这几个网址返回不同的内容。

除了上面因为带与不带 www 造成的，以及结尾是否带 index.html 文件名造成的不规范网址，网址规范化问题还可能由于如下原因出现：

- CMS 系统原因，使同一篇文章（也可以是产品、帖子等）可以通过几种不同的 URL 访问。
- URL 静态化设置错误，同一篇文章中有多个静态化 URL。
- URL 静态化后，静态和动态 URL 共存，都有链接，也都可以访问。
- 目录后带与不带斜杠。

http://www.domainname.com 和 http://www.domainname.com/是不同网址，但其实是一个页面。

- 加密网址。

http://www.domainname.com 和 https://www.domainname.com 同时存在，都可以访问。

- URL 中有端口号。

http://www.domainname.com:80 和 http://www.domainname.com。

- 跟踪代码。有的联署计划或广告服务在 URL 后面加跟踪代码，如 http://www.domainname.com/?affid=100
- 显示的都是首页内容。

4.10.2 网址规范化问题

网站出现多个不规范网址会给搜索引擎收录和排名带来很多麻烦。比如，网站首页应该是固定的，只有一个，但很多站长在链接回首页时所使用的 URL 并不是唯一的，一会儿连到 http://www.domainname.com，一会儿连到 http://www.domainname.com/index.html。

虽然不会给用户造成什么麻烦，因为这些网址其实是同一个文件，但是给搜索引擎造成了困惑，哪一个网址是真正的首页呢？哪一个网址应该被当做首页返回呢？

如果网站上不同版本的网址同时出现，那么两个或更多版本的 URL 都可能被搜索引擎收录，这就会造成复制内容。搜索引擎计算排名时必须找到所谓规范化的网址，也就是搜索引擎认为的最合适的 URL 版本。

网址规范化造成几个问题：

- CMS 系统在不同地方链接到不同的 URL，分散了页面权重，不利于排名。
- 外部链接也可能指向不同 URL，分散权重。
- 搜索引擎判断的规范化网址不是站长想要的那个网址。
- 如果网址规范化问题太严重，也可能影响收录。一个权重不很高的域名，能收录的总页面数和蜘蛛总爬行时间是有限的。搜索引擎把资源花在收录不规范的网址上，留下给独特内容的资源就减少了。
- 复制内容过多，搜索引擎可能认为有作弊嫌疑。

4.10.3 解决网址规范化问题

要解决 URL 规范化问题有很多方法，比如：

- 确保使用的 CMS 系统只产生规范化网址，无论是否有静态化。
- 所有内部链接保持统一，都指向规范化网址。以带与不带 www 为例，确定一个版本为规范化网址，网站内链接统一使用这个版本。这样搜索引擎也就明白哪一个是站长希望的规范化网址。由于用户习惯，通常选择带 www 的版本为规范化网址。
- 在 Google 管理员工具中设置首选域。
- 使用 3 01 转向，把不规范化 URL 全部转向到规范化 URL。下一节将深入讨论。
- 使用 canonical 标签。后面将深入讨论。

- 提交给搜索引擎的 XML 网站地图中全部使用规范化网址。

但这些方法都各有局限：

- Google 管理员工具不适用于其他搜索引擎。
- 有的网站因为技术原因做不了 301 转向。
- CMS 系统经常不受自己控制。
- 内部链接自己可以控制，但外部链接不受控制。

所以，虽然有解决方法备选，但到目前为止网址规范化一直是困扰站长及搜索引擎的一个问题。据估计，网上有 10%～30%的 URL 是内容相同但 URL 不一样的不规范化网址。为了万无一失，经常需要综合使用多个方法。

4.10.4 301 转向

1．什么是 301 转向

301 转向（或叫 301 重定向、301 跳转）是用户或蜘蛛向网站服务器发出访问请求时，服务器返回的 HTTP 数据流中头信息（header）部分状态码的一种，表示本网址永久性转移到另一个地址。

其他常见的状态码还包括：

- 200 表示一切正常
- 404 网页不存在
- 302 临时性转向
- 500 内部程序错误

网址转向还有其他方法，如 302 转向、JavaScript 转向、PHP/ASP/CGI 程序转向、Meta Refresh 等。除了 301 转向外，其他方法都是常用的作弊手法。虽然方法本身没有对错之分，但被作弊者用多了，搜索引擎对可疑的转向都很敏感。

2．301 转向传递权重

网页 A 用 301 重定向转到网页 B，搜索引擎可以肯定网页 A 永久性改变地址，或者说实际上不存在了，搜索引擎就会把网页 B 当做唯一有效目标。这是搜索引擎唯一推荐的不会产生怀疑的转向方法，更重要的是，网页 A 积累的页面权重将被传到网页 B。

所以，假定 http://www.domainname.com 是选定的规范化网址，下面几个网址：

- http://domainname.com
- http://www.domainname.com/index.html
- http://domainname.com/index.html

都做301转向到http://www.domainname.com，搜索引擎就知道http://www.domainname.com是规范化网址，而且会把上面列的网址权重传递集中到规范化网址。

目前Google会传递大部分权重，但不是百分之百权重。百度会传递所有权重。Google对301的识别、反应、完成权重传递，需要1～3个月时间。百度对301处理则比较保守，新旧URL会同时存在于数据库很长时间，经常看到做了301转向一年的URL还出现在百度结果中。

3．怎样做301转向

如果网站使用LAMP（Linux+Apache+MySQL+PHP）主机，可以使用.htaccess 文件做301转向。.htaccess 是一个普通文字文件，用Notepad等文字编辑软件创建和编辑，存在网站根目录下。.htaccess 文件中的指令用于目录特定操作，如密码保护、转向、错误处理等。

比如把页面/old.htm 301转向到http://www.domain.com/new.htm，可以在.htaccess文件中放上这个指令：

redirect 301 /old.htm http://www.domain.com/new.htm 或 redirect permanent /old.htm http://www.domain.com/new.htm

把所有不带 www（http://domain.com）版本的网址 301 转向到带 www 的版本（http://www.domain.com），包括：

http://domain.com/about.htm 转到 http://www.domain.com/about.htm 及 http://domain.com/dir/index.htm转到http://www.domain.com/dir/index.htm等，还要用到mod_rewrite模块，.htaccess文件指令是：

```
Options +FollowSymLinks
RewriteEngine on
RewriteCond %{HTTP_HOST} ^domain.com [NC]
RewriteRule ^(.*)$ http://www.domain.com/$1 [L,R=301]
```

如果网站用的是Windows主机，可以在控制面板做301转向设定。

纯静态HTML页面无法做301转向。在HTML里能做JavaScript或Meta Refresh转向，但不能做301转向。HTML一被读取，就已经返回200 Ok状态码了。

如果页面是ASP或PHP，还可以做301转向。

ASP代码：

```
<%@ Language=VBScript %>
<%
Response.Status="301 Moved Permanently" Response.AddHeader "Location", "
http://www.domain.com"
>
```

PHP 代码：

```
Header( "HTTP/1.1 301 Moved Permanently" );
Header( "Location: http://www.domain.com" );
?>
```

4．301 转向的其他用途

除了解决网址规范化问题，还有很多需要做 301 转向的情形。比如，为保护版权，公司拥有不同 TLD 的多个域名：

- company.com
- company.net
- company.org
- company.com.cn
- company.cn

为避免造成大量复制内容，应该选定一个为主域名，如 company.com，其他域名做 301 转向到 company.com。

或者公司注册了全称域名 longcompanyname.com，但太长，也注册了缩写域名方便用户记住 lcn.com，其中一个做主域名，另一个 301 转向到主域名。

网站改版也经常需要用到 301 转向，如页面删除、改变地址、URL 命名系统改变等。

动态 URL 静态化也可能要做 301，将旧的、动态的 URL 做 301 到新的、静态的 URL。

4.10.5 Canonical 标签

2009 年 2 月，Google、雅虎、微软共同发布了一个新的标签 canonical tag，用于解决网址规范化问题。

简单说，就是在 HTML 文件的头部加上这样一段代码：

```
<link rel="canonical" href="http://www.example.com/product.php?item=
swedish-fish" />
```

意义就是这个网页的规范化网址应该是：

http://www.example.com/product.php?item=swedish-fish

下面这些 URL 都可以加上这段 canonical 标签：

- http://www.example.com/product.php?item=swedish-fish&category=gummy-candy

- http://www.example.com/product.php?item=swedish-fish&trackingid=1234&sessionid=5678

这些 URL 的规范化网址就都成为：

http://www.example.com/product.php?item=swedish-fish

Canonical 标签相当于一个页面内的 301 转向，区别在于用户并不被转向，还是停留在原网址上，而搜索引擎会把它当做是 301 转向处理，把页面权重集中到标签中指明的规范化网址上。

⚠ 另外有几个细节需要注意：

- 这个标签只是一种建议或暗示，而不是指令，它不像 robots 文件那样是个指令。所以搜索引擎会在很大程度上考虑这个标签，但并不是百分之百。搜索引擎还会考虑其他情况来判断规范化网址。这也防止站长有可能把网址弄错。
- 标签既可以使用绝对地址，也可以使用相对地址。通常使用绝对地址比较保险。
- 指定的规范化网址上的内容，与其他使用这个标签的非规范化网址内容可以有一些不同，不一定完全一样。比如电子商务网站上多按价格、时间、尺寸升降排序，生成的 URL 全都不一样，但内容大体相同，只有细小区别，就可以使用这个标签。
- 指定的规范化网址可以是不存在页面，返回 404，也可以是还没有被收录的页面。但是不建议这么做，别没事找事。
- 这个标签可以用于不同域名之间。

有些网站由于技术限制不能做 301 转向，canonical 标签就显得非常灵活，不需要任何特殊服务器组件或功能，直接写在页面 HTML 中就可以了。

目前 Google、雅虎、必应都支持 canonical 标签，不过雅虎和必应收录页面后做出反应的时间非常长。百度还不支持 canonical 标签。

4.11 复制内容

复制内容也可以称为重复内容。复制内容指的是两个或多个 URL 内容相同，或非常相似。复制内容既可能发生在同一个网站内，也可能发生在不同网站上。

4.11.1 产生复制内容的原因

下面这些原因可能造成复制内容：

前面讨论的网址规范化问题会产生复制内容。

- 代理商和零售商从产品生产商那里转载产品信息。这倒没什么不对，一般生产商也都同意，没有版权问题。但是绝大部分代理商、零售商、批发商都是直接 copy，而不做任何改动，大家用的都是一模一样的产品说明，所以这些电子商务网站上充斥着大量复制内容。
- 打印版本。很多网站除了正常供浏览的页面外，还提供更适于打印的页面版本，如果没有用适当方式禁止抓取，这些打印版本网页就会变成复制内容。
- 网站结构造成的各种页面版本。如产品列表按价格、评论、上架时间等排序页面，博客的分类存档、时间存档等。
- 网页内容由 RSS 生成。有很多网站，尤其是新闻类网站，用其他网站的 RSS feed 生成网站内容，这些内容在原始出处和很多其他类似网站上都已经出现过很多次了。
- 使用 Session ID。搜索引擎蜘蛛在不同时间访问网页的时候，被给予了不同的 Session ID，实际上网页内容是一样的。由于 Session ID 参数不同，就被当成了不同的网页。
- 网页实质内容太少。每个网页上都不可避免地有通用部分，比如导航条、版权声明、广告等。如果网页的正文部分太短，内容数量还抵不上通用部分，就有可能被认为是复制内容页面。
- 转载及抄袭。有时是其他人抄袭了你的网站内容，有时是善意的转载，有时是作者自己在不同网站发布文章，这些都会造成复制内容。
- 镜像网站。镜象网站曾经很流行，当一个网站太忙太慢的时候，用户可以通过替代镜像来看内容或下载，这也有造成复制内容的风险。
- 产品或服务类型之间的区别比较小。比如有的网站把自己的服务按地区进行分类，但实际上提供给每个地区的产品或服务都是一样的。这些按地区分类的页面只是把地名改了改，其他服务内容说明全都一样。
- URL 任意加字符还是返回 200 状态码。有的网站由于技术原因，用户在 URL 后加上任意字符或参数，服务器还能正常返回 200 状态码，并返回没加上任意字符时一样的重复内容页面。

检查页面是否有复制版本相对简单。拿出页面正文中的一句话，加上双引号，在搜索引擎中搜索一下，从结果中就能看到是否有多个页面包含这句话。一般来说，随机挑选的一个句子，完整出现在另一篇无关文章中的可能性很低。

举一个比较极端却很清楚的例子。我为了试验搜索引擎是否使用关键词标签，曾经在一篇博客帖子中提到“伍疗踢瓜 sdfghj”这么一个字符串。在我发这个帖子之前，搜索引擎没有这个字符串的任何结果。现在搜索一下，可以清楚看到有不少因为转载或抄袭形成的复制内容，如图 4-9 所示。

在 Google 搜索得到的结果就更多了，将近 1000 个页面。令人无奈的是，绝大部分转载没有按版权声明链接到原出处，更有很多直接连作者都给改了。

图 4-9　转载、抄袭造成的复制内容

4.11.2　复制内容的害处

很多 SEO 对复制内容有个误解，认为网站上有复制内容，搜索引擎就会惩罚。其实搜索引擎并不会因为网站有少量复制内容而惩罚或降权。搜索引擎做的只是从多个页面中尽量挑选出真正的原创版本给予应有的排名，其他复制版本不在搜索结果中返回或排在比较靠后的位置。

不过既然交给搜索引擎去判断，就有可能判断失误，把本来是原创的页面当做转载或抄袭的复制内容，因而排名不好。这对原创页面来说确实就像是惩罚，不过不是因为复制内容本身，而是因为搜索引擎判断原创失败。

同站出现复制内容看似没有那么严重，不管搜索引擎判断哪个版本是原创，都是自己网站上的页面。问题是，搜索引擎认为的最合适的 URL 与站长自己认为的最合适的 URL 有可能不一样，站长优化和做链接时把精力放在页面 A，搜索引擎却认为页面 B 最好（A 与 B 内容一样），站长花在页面 A 上的精力就浪费了很多。

同一个网站内的复制内容会分散权重。既然页面在网站上出现，就必然有链接连向这些页面。如果一篇独特内容只出现在一个网址，网站上的链接就能集中到这一个网址，使其排名能力提高。链接分散到多个网址，会使得每一个网址排名能力都不突出。外部链接也同样，很可能分散到不同 URL。

同站复制也造成收录过多没有意义的页面，在域名权重不高的情况下，可能挤占了其他独特内容的收录机会。

如果网站上存在大量复制内容，可能使搜索引擎对网站质量产生怀疑，导致惩罚。

4.11.3　消除复制内容

网址规范化问题造成的复制内容，前面已经讨论过。最好的解决办法是，确保一篇

文章只对应一个 URL，不要出现多个版本，网站所有内部链接统一连到这个 URL。某些时候需要使用 301 转向。

不是由于网址规范化造成的同站内复制内容，解决的最好方法是选取一个版本允许收录，其他版本禁止搜索引擎抓取或收录。既可以使用 robots 文件禁止抓取，也可以使用 noindex meta robots 标签禁止索引。连向不希望收录的复制内容的链接使用 NoFollow，JavaScript 等阻止蜘蛛爬行。

另一个解决复制内容的方法是使用 canonical 标签。canonical 标签既可以应用在网址规范化引起的复制内容上，也可以用在其他情况下。比如网站上同一个款式的鞋子可能分为不同型号，型号之间唯一的区别只是颜色。这些型号生成多个网址，产品说明也几乎完全一样，只是说明颜色的地方不一样。这时就可以使用 canonical 标签，用户不会被转向，还是看到不同页面，但搜索引擎会把权重集中到其中一个型号上，从而避免复制内容。

带有 Session ID 的页面也可以使用 canonical 标签，如页面 http://www.example.com/page-a.html 放上代码：

```
<link rel="canonical" href="http://www.example.com/page-a.html" />
```

这样，后面无论生成什么 Session ID：

```
http://www.example.com/page-a.html?sessionid=123456
http://www.example.com/page-a.html?sessionid=456789
```

上面的 URL 都会被搜索引擎把权重集中到 http://www.example.com/page-a.html。

不同网站之间复制内容解决起来就比较麻烦，因为其他网站上的内容是无法控制的。能够做的只有两点。一是在页面中加入版权声明，要求转载的网站保留版权声明及指向原出处的链接。有些聚合网站会抓取 feed 自动生成内容，所以在 feed 中也要加入版权声明和链接。一般来说，原创版本的外部链接会比转载多一些，就算在中文互联网这个对版权极不重视的环境下，也还会有一些站长转载时保留原出处链接。对搜索引擎来说，指向原出处的链接是判断原创的最重要信号。

另一点就是坚持原创，假以时日必定能够增加网站权重，使网站上内容被判断为原创的机会增加。

如果有其他网站大量抄袭，造成原创内容不能获得排名，也可以考虑联系对方，要求加上版权链接或删除抄袭内容，或者向对方主机、域名提供商投诉，向搜索引擎投诉等。

4.12 绝对路径和相对路径

绝对路径指的是包含域名的完整网址。相对路径指的是不包含域名的、被链接页面

相对于当前页面的相对网址。

比如页面 A 的 URL 是：

http://www.domain.com/pageA.html

页面 B 的 URL 是：

http://www.domain.com/pageB.html

页面 A 链接到页面 B 时使用这种代码：

```
<a href="../pageB.html">
```

就是相对路径。

如果使用完整的 URL：

```
<a href="http://www.domain.com/pageB.html">
```

也就是浏览器地址栏中所显示的完整 URL，就是绝对路径。

网站应该使用绝对路径还是相对路径问题？对 SEO 有什么影响？没有绝对答案，两者各有优缺点。

4.12.1 绝对路径

绝对路径 URL 优点是：

- 如果有人抄袭、采集你的网站内容，抄袭者比较懒，连页面里面的链接一起原封不动抄过去，绝对路径链接还会指向你的网站，增加网站外链及权重。
- 网站有 RSS 输出时，内容会被一些 Feed 聚合网站抓取显示。同样，页面里指向原网站的链接会被保留。
- 有助于预防和解决网址规范化问题。假设站长希望被收录的 URL 是带 www 的版本，由于技术原因不能从 http://domain.com 做 301 转向到 http://www.domain.com。所有页面中的链接使用绝对路径如：

http://www.domain.com/article.html

http://www.domain.com

链接绝对路径硬编码入 HTML 文件，这样就算有蜘蛛或用户偶然访问不带 www 的版本，如 http://domain.com/article.html，这个页面上的绝对路径链接还是会把蜘蛛和用户带回到 www 的 URL 版本，有助于搜索引擎蜘蛛识别到底哪个版本是规范化的。

就算网页移动位置，里面的链接还是指向正确 URL。

绝对路径缺点是：

- 除非链接是动态插入的，不然不好在测试服务器上进行测试。因为里面的链接将直接指向真正域名的 URL，而不是测试服务器上的 URL。
- 除非链接是动态插入的，不然移动页面将比较困难。因为页面位置发生变化，其他页面连向本页面的链接却可能无法跟着变化，还指向原来的已经硬编码的绝对路径。
- 代码比较多。链接数量大时，多出来的字符可能使 HTML 文件变大不少。

4.12.2 相对路径

相对路径正好相反。

优点是：

- 移动内容比较容易，不用更新其他页面上的链接。
- 在测试服务器上进行测试也比较容易。
- 节省代码。

缺点是：

- 页面移动位置，里面的链接可能也需要改动。
- 被抄袭和采集对网站没有任何益处。不过很多采集软件其实是可以自动鉴别绝对路径和相对路径的，所以使用绝对路径有助于自己的链接也被抄到采集网站上，只在某些情况下是有效的。
- 搜索引擎解析 URL 时可能出错，不能正确读取页面上的链接 URL。

如果不能做 301 转向，因而产生了严重的网址规范化问题时，使用绝对路径有助于解决网址规范化问题。如果文章被大量转载、抄袭，使用绝对路径可以带来一些外部链接。除此之外，使用相对路径比较简单。在正常情况下，相对路径不会对网站有什么副作用，绝对路径也不会有什么特殊好处。搜索引擎错误解析相对路径 URL 的可能性是非常低的。

在正确解析 URL 的前提下，绝对路径和相对路径本身对排名没有任何影响。

4.13 网站地图

网站无论大小，单独的网站地图页面都是必需的。通过网站地图，不仅用户可以对网站的所有内容一目了然，搜索引擎也可以跟踪链接爬行到网站所有主要部分。

4.13.1 HTML 网站地图

网站地图有两种形式。第一种称为 HTML 版本的网站地图，英文是 sitemap，s 需小写，特指 HTML 版网站地图。HTML 版本网站地图就是用户可以在网站上看到的、列出网站上所有主要页面链接的页面。

对一些主导航必须使用 JS 脚本或 Flash 的网站（虽然我想不到什么原因必须要这么做），网站地图是搜索引擎找到网站所有页面的重要补充途径。

对小网站来说，网站地图页面甚至可以列出整个网站的所有页面。

但对稍具规模的网站来说，一个网站地图页面不可能罗列所有页面链接，可以采取两种办法。一种方法是网站地图只列出网站最主要部分链接，如一级分类、二级分类。另一种方法是将网站地图分成多个文件，主网站地图列出通往次级网站地图的链接，次级网站地图再列出一部分页面链接。多个网站地图页面加在一起，可以列出所有或绝大部分重要页面。

实际上一个具有良好导航系统和链接结构的网站，并不一定需要完整的、列出所有页面的网站地图，因为网站地图与分类结构经常是大同小异的。

4.13.2 XML 网站地图

网站地图的第二种形式是 XML 版本的网站地图。英文 Sitemap 中的 S 大写时通常特指 XML 版本的网站地图。

XML 网站地图由 Google 于 2005 年首先提出，2006 年微软、雅虎都宣布支持。2007 年各主要搜索引擎都开始支持 robots 文件指定网站地图位置。百度现在也支持 XML 版本网站地图。

XML 版本网站地图由 XML 标签组成，文件本身必须是 utf8 编码。网站地图文件实际上就是列出网站需要被收录的页面 URL。最简单的网站地图可以是一个纯文本文件，文件只要列出页面 URL，一行列一个 URL，搜索引擎就能抓取并理解文件内容。

标准版本的 XML 网站地图文件如下列代码所示：

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
   <url>
      <loc>http://www.example.com/</loc>
      <lastmod>2010-01-01</lastmod>
      <changefreq>monthly</changefreq>
      <priority>0.8</priority>
   </url>
</urlset>
```

其中 urlset 标签是必需的，声明文件所使用的 Sitemap 协议版本。

url 标签也是必需的，是它下面所有网址的母标签。

loc 标签也是必需的，这一行列出的就是页面完整 URL。

lastmod 是可选标签，表示页面最后一次更新时间。

changefreq 是可选标签，代表文件更新频率。标签值包括：

- always，一直变动，指的是每次访问页面内容都不同。
- hourly，每小时。
- daily，每天。
- weekly，每星期。
- monthly，每月。
- yearly，每年。
- never，从不改变。

网站地图中声明的更新频率对搜索引擎来说只是一个提示，供搜索引擎蜘蛛参考，但搜索引擎不一定真的认为页面更新频率就是站长自己声明的那样。

priority 是可选标签，表示 URL 的相对重要程度。可选 0.0 到 1.0 之间的数值，1.0 为最重要，0.0 为最不重要。默认重要程度值为 0.5。站长可以使用 priority 标签告诉搜索引擎这个 URL 的优先级，比如通常首页肯定是 1.0，分类页面可能是 0.8，其他更深层页面重要性依次下降。这里所标志的重要程度只是相对于这个网站内部的 URL 所说的，与其他网站的 URL 重要性无关。所以把页面重要性标为 1.0，并不能让搜索引擎认为这个页面比其他网站的页面更重要，只是告诉搜索引擎这个页面在本网站内是最重要的。

XML 网站地图文件最多可以列出 5 万个 URL，文件不能超过 10MB。如果网站需要收录的 URL 超过 5 万个，可以使用网站地图索引文件，也就是分成多个网站地图文件。一个网站最多可以有 1000 个 XML 网站地图文件，也就是可以列出 5000 万个网址。

制作好网站地图文件后，可以有两种方式通知搜索引擎网站地图的位置，一是在站长工具后台提交网站地图文件。目前 Google、雅虎、必应都已经开通站长工具，所有站长都可以注册。2010 年 4 月份，百度也推出站长工具，正在内测中，还未对所有站长开放。百度站长工具的主要功能之一也是提交网站地图。

另外一种方式是在 robots.txt 文件中通知搜索引擎网站地图文件位置，代码如下所示：

```
Sitemap: http://www.example.com/sitemap.xml
```

所有主流搜索引擎，包括百度、Google、雅虎、必应，都支持 robots 文件指定网站地图文件位置。

通过 XML 网站地图通知搜索引擎要收录的页面，只能让搜索引擎知道这些页面的存在，并不能保证一定被收录，搜索引擎还要看这些页面的权重是否达到收录的最低标准。所以 XML 网站地图只是辅助方法，不能代替良好的网站结构。

大中型网站提交网站地图通常有比较好的效果，能使收录增长不少。但也有的网站，尤其是小型网站，提交网站地图没有什么效果，有的站长甚至认为有反效果。

4.14 内部链接及权重分配

前面提到网站结构优化要解决的最重要的问题包括收录及页面权重分配。在理想情况下，经典树形结构应该是比较好的链接及权重分配模式。但是由于不同网站采取的技术不同，要实现的功能、网站目标、重点要解决的营销问题都很可能不同，某些看似是树形结构的网站，仔细研究起来其实是奇形怪状的。每个网站有自己的特殊问题需要解决，没有可以适用于所有网站的结构优化秘诀，必须具体问题具体分析。为了扩展读者的思路，本节列举一些可能会遇到的情况及解决方法。

4.14.1 重点内页

首先考虑一个最常见也最简单的情况。一般来说，网站首页获得的内外部链接最多，权重最高。首页链接到一级分类页面，这些一级分类页面权重仅次于首页。大部分网站有多层分类，权重依次下降，权重最低的是最终产品页面。

有时某些具体产品页面却需要比较高的权重，比如某些转化率最高、利润率最高或者新推出的重点产品，搜索次数很多的产品（有的产品名称搜索次数相当大，超过一些分类名称），还有为特定节日或促销活动制作的专题页面。这些页面按照经典树形结构安排，离首页通常有一定距离，权重不会太高。要想使这种重点内页获得高权重，最简单的方法就是在首页上直接加上几个重点内页的链接，甚至可以在侧栏推荐、促销部分加上全站链接。

在很多电子商务网站上，首页展示的是最新产品、热门产品等。这些产品的选择其实是有学问的，并不一定真的按发布时间列出最新产品，或真的是产生订单最多的产品。站长可以放上自己想重点推广的产品页面，使这些内页对应的产品名称排名提高。

依我的经验，把产品内页链接放在首页上，哪怕没有其他外部链接的支持，这些内页的排名也会有显著提高。

4.14.2 非必要页面

每个网站都有一些在功能及用户体验方面很必要、但在 SEO 角度没必要的页面，比如隐私权政策、用户登录页面、用户注册页面、联系我们、甚至还包括“关于我们”页

面。从用户角度看，这些页面是必需的功能或有助于提高网站信任度。不过搜索引擎既不能填表注册，也不能登录，一般网站也不会想优化“隐私权政策”这种关键词，这些页面既没有必要也不太可能获得任何排名。

麻烦的是这些页面通常在整个网站每个页面上都会有链接，它们的权重将仅次于首页，与一级分类页面相似，可能更高。不得不说这是一种权重浪费。为降低这些非必要页面的权重，一般可以采取几种方式。

（1）只在首页显示链接，其他页面干脆取消链接，如隐私权政策，关于我们等页面。

（2）使这些页面的链接不能被跟踪或传递权重，例如使用 NoFollow 标签或使用 JavaScript 链接。某些必须在所有页面显示的链接可以这样处理，如用户注册及登录页面。

还可以将这些页面的链接通过一个程序页面做转向，如链接做成 redirect.php?page.html，redirect.php 程序只实现一个功能，就是根据后面的参数自动转向到 page.html，然后用 robots 文件禁止爬行 redirect.php 这个文件。

除了上面提到的几种明显的非必要页面，很多网站其实存在更多非必要页面。比如电子商务网站列出的帮助信息、购物付款流程、送货信息、公司新闻等，这些页面从 SEO 角度看都没有什么价值。

SEO 人员应该对网站所有版块了如指掌。凡是在产品分类及具体产品页面之外的信息，都要问问自己，这些页面站在 SEO 角度看是必需的吗？能优化什么关键词？尽量减少能够传递权重的全站链接到非必要页面。

4.14.3 大二级分类

典型树形结构首页链接到一级分类，一级分类页面再列出二级分类，这样，只要二级分类数目相差不太悬殊，权重值在二级分类页面上是大致平均分配的。不过有的时候某些二级分类下的产品数远远多于其他小一些的二级分类，甚至产品太多的二级分类下还可能再列三级分类。平均分配权重的结果就是，小分类充分收录，产品数量大的大分类有很多产品页面因为权重稀释无法收录。

要解决这个问题，思路就是提高大二级分类页面的权重，使它能带动的产品页面增多。现在很多网站导航系统采用 CSS 下拉菜单方式，鼠标放在一级分类链接上时，菜单向右或向下拉出一部分二级分类页面。当由于空间有限和用户体验因素不能显示所有二级分类时，选择显示哪些二级分类就是个学问。显然不能按拼音或字母顺序把排在前面的 5 个或 10 个放入主导航。一种方法是从用户体验出发，先选择热门二级分类。再一种方法是选择包含产品数量最多、需要权重支持才能充分收录的二级分类。这两者有时候是重合的，热门分类也是产品最多的分类，有时候不尽然。

这个原则同样适用于多层分类。如果网站有三层分类页面，应该计算出每个三级分类下有多少产品，想办法把这些大三级分类页面放在首页上，如果可能，放在尽量多的导航中。

4.14.4 翻页过多

稍大型的商务或信息类网站都可能会在产品列表，也就是最末一级的分类页面上，存在翻页过多的问题。通常产品列表会显示 10 个或 20 个产品，然后列出翻页链接，除了“上一页”和“下一页”，网站可能列出 5 个、10 个翻页链接。用户点击页面 10，第 10 页上又会列出 10～19 页的翻页，如图 4-10 所示。

图 4-10　常见的翻页系统

可以简单计算一下，如果这个分类下有 1000 个产品，每个页面列出 20 个产品，就需要 50 个页面才能显示完所有产品。如果页面列 10 个翻页链接，那么第 50 个页面上的产品就需要从第一个产品列表页面点击 4 次才能到达，再加上分类页面本身与首页的距离，第 50 个页面上的产品距首页可能有七八次点击的距离了。

如果像图 4-10 那样列出 5 个翻页，当前页面居中，要到达第 50 页已经是二三十次点击之后了。

很多网站在某个分类下有成千上万的产品，可以想象，按传统翻页导航，列在后面的产品可能需要点击几十上百次才能到达。如果没有适当的结构优化，这些产品页面被收录的可能性几乎为零。

解决这个问题的最佳方式是再次分类。假设一个分类下有 2000 个产品，排在最后的页面按上面同样的情况，需要点击 10 次才能到达。如果把这个分类页面再次细分为 20 个子类，那么每一个产品页面就都在两次点击距离之内。

稍微计算一下就会知道，多一层分类给大中型网站带来的结构利益是巨大的。我们假设一级产品分类由于用户体验原因只能分 10 个，每个一级分类下面都能再列出 30 个二级分类，这样二级分类总数就能达到 300。每个二级分类下最多有 200 个产品（每页 20 个产品，10 页显示完所有产品），就能保证每个产品页面都在距首页 4 次点击之内，总共能带动的产品页面数为 6 万个。而如果加多一级分类，每个二级分类下再分 30 个三级分类，总三级分类数目将达到 9000 个。如果每个三级分类下有 200 个产品，所能带动的总产品数就达到了 180 万，而到达每个产品页面点击数只增加了一次，这比列出几十几百个翻页链接要好得多。

另外一个解决方式是对翻页链接进行格式变化，比如将翻页链接改为下面这种格式：

1，2 ，3，4，5，10，20，30，40，50

这样排在第 50 页的产品，只要再多一次点击也能达到。

如果产品数再多，甚至可以把翻页做成两排：

1，2，3，4，5，6，7，8，9，10

20，30，40，50，60，70，80，90，100

如上面所示两排翻页结构，2000 个产品多一次点击就可以全部达到。页面可以列出多少翻页链接及链接之间的步长，SEO 可以根据产品数量和网站权重调节。

4.14.5 单一入口还是多入口

一般来说，网站的首页和分类页面收录不会有什么问题，除非主导航系统有严重蜘蛛陷阱，或者网站已经被惩罚。大部分网站在结构方面面对的挑战，是使更多最终产品页面被收录。就算尽量把网站结构扁平化，当产品数量巨大时，实际上已经不太可能足够扁平。在这种情况下，要让最终产品页面被收录，有两个策略：一是多入口，二是单一入口，在选择上需要谨慎。

多入口指的是通向最终产品页面的链接路径有多条。比如典型电子商务网站的产品页面，一定会出现在相应的分类产品列表中，还可能出现在不同的排序页面上（按价格排序、按热门程度排序、按上架时间排序、按评论数排序等），以及不同的显示方式（按栅格显示、列表显示），也可能出现在相关的品牌或生产商产品列表中，也可能出现在搜索页面或标签聚合页面中，如图 4-11 所示。

图 4-11　多种显示、排序方式

再比如博客系统，同一篇帖子除了在博客主列表中出现，还会出现在分类存档、按时间存档、标签页面，有的还会出现在作者分类页面。其他 CMS 系统也大多具备这种多入口结构通向最终产品页面。

这种结构的优势是为最终页面提供了多条爬行收录渠道，东方不亮西方亮。由于某种原因没从分类页面爬行，还可能从其他页面爬行抓取。提供的入口越多，被收录的机会越大。

缺点是这些入口页面本身也占用网站收录页面名额，而且往往造成很多相似内容。一个给定的网站，权重是大致固定的，所能收录的总页面数也是有上限的。收录的分类页面、各种排序页面、品牌生产商页面、搜索标签页面越多，给最终产品剩下的收录名

额就越少。要提高整个网站的权重，唯一的办法就是增加外部链接。如果网站权重足够高，能带动的收录页面数远远大于实际页面数，提供多入口就是最佳方式，因为那些冗余的入口页面并不会挤占产品页面的名额。

但是如果网站权重比较低，产品数又很大，就可能需要使用单一入口方法，也就是从首页到产品页面只提供单一通路，通常也就是主导航的分类页面，其他各种排序页面、品牌、生产商页面，全部使用 JavaScript 脚本或 NoFollow 标签，甚至 robots 文件，阻挡搜索引擎收录。对某些网站来说，多入口页面本身数量就很巨大，会占用很多收录页面份额。只要网站分类系统、导航及翻页设计合理，提供单一入口也可以达到收录尽量多最终产品页面的目的。

哪种方法最适合还得看网站自身情况，如域名权重、实际总页面数等。

4.14.6 相关产品链接

无论单一入口还是多入口结构，对最终产品页面来说都可能有一个缺陷，那就是太过规则，有时候会造成某个部分的产品页面都不能被收录。单一入口结构更明显。比如如果某个分类页面因为导航设计不合理，离首页太远没有被收录，这个分类下的所有产品就都无法被收录。再比如博客系统中，发表比较早的帖子，无论从哪个入口渠道看，都会被推到网站更深层，离首页比较远，老帖子虽然一般没有收录问题（早就被收录了），但权重却随着时间推移而下降。

在产品页面生成相关产品链接，可以在一定程度上解决这个问题。这里所说的相关产品链接，不是写文章或发布产品信息时人工在正文中加进去的链接，而是系统通过某种机制自动生成的、连向其他产品页面的链接。

好的相关产品链接应该具有比较强的随机性，与正常的分类入口区别越大越好。常见的相关产品链接生成方法包括：

（1）购买这个产品的用户还购买了哪些其他产品。这种链接通常不会是同时上架、产品序号相连的页面，用户购买过的产品之间不一定有什么联系，往往横跨不同分类、品牌。

（2）同一个品牌或生产商的其他产品。同一个生产商或品牌，常常有不同分类下的产品，最终产品页面列出同一个生产商提供的不同分类的产品链接，也为更多产品提供了较为随机的入口。

（3）由标签生成的相似产品。标签由站长人工填写，或程序自动提取关键词，得到的标签与分类名称并不相同。通过标签聚合相似产品，也具有比较大的随机性。

（4）最简单的相关文章链接，就是在博客和新闻类网站中常看到的“上一篇”和“下一篇”这种链接。不过这样的相关链接对最终页面收录意义不大，因为时间上前后相连的文章本来就在时间存档及分类页面中相连，被同时收录和同时不收录的概率更大。与此类似，有的 CMS 系统在产品页面列出这个产品之前和之后的几个产品，意义也不大。

总之，相关产品链接要尽量随机，使本来不相连的页面能够交叉链接起来，为某些通过正常分类结构无法达到的区域提供入口。

4.14.7 锚文字分布及变化

前面提到过，合理的网站结构是在网站中分配锚文字的重要方法之一。最灵活常见的是在页面正文中或人工或自动加上其他页面的内部链接，链接锚文字可以有各种选择。这方面的应用学习目标非维基百科莫属。

相比之下，网站导航系统中锚文字的分布及变化很少有人注意。因为导航系统名称相对固定，分类该叫什么名称就叫什么名称，绝大部分网站在全站导航中不会给分类链接锚文字做任何变化。

其实仔细研究一下，即使在导航系统中锚文字也可以有变化。比如顶部导航使用“电脑”这个词，左侧导航改成“计算机”，或者左侧导航使用“快速减肥”作为分类链接锚文字，在面包屑导航中同样的分类改成“迅速减肥”。很少用户会注意到这种极细微的差别，就算注意到也无伤大雅，因为意义完全相同，对用户浏览网站没有任何影响。对分类页面来说，却可以增加不同的导入链接锚文字。

如果分类页面可以有更多具有相同意义的名称，还可以在导航系统中找到更多可以变化的地方。比如在网站不同部分（分类首页及其下所有产品页面），导航系统使用的锚文字也可以不同。如电脑外设部分页面连向移动硬盘分类就用“移动硬盘”做锚文字，在电脑软件部分所有页面指向同一个分类（移动硬盘）时，锚文字可以改成“便携式硬盘”，在电脑耗材所有页面上，导航系统锚文字又可以换成“USB 硬盘”。

当然这里只是举例，具体网站是否需要在导航系统中变化锚文字，该选择哪些锚文字，应该先做关键词研究，再做决定。前提是，用做锚文字变化的词意义必须一样，不能影响用户体验，而且搜索次数差不多，都需要锚文字加强相关性。

4.14.8 首页链接 NoFollow

很多页面上会有多个链接连向同一个 URL，比如几乎网站的每一个页面上都有多个链接连向首页，顶部 logo、顶部导航、左侧导航、页脚、版权声明等处，都可以有链接到首页。很多人的观察和实验表明，当页面上出现多个链接到同一个网址时，第一次出现的链接最重要，第一个链接的锚文字也最重要。

如果页面上出现的第一个首页链接是顶部 logo，那么 logo 图片的 ALT 文字就相当于锚文字，需要放上首页的目标关键词。

也有的人认为图片 ALT 文字比真正的文字链接锚文字作用要小，因而不愿意把最重要的第一个链接放在图片上。在不影响用户体验的情况下，网站顶部 logo 也可以不放链接。这时页面上连向首页的第一个链接，往往就是顶部导航最左侧的首页链接。可惜的是，这里的链接锚文字一般来说就是“首页”两个字，而不能加上关键词。有人曾经尝试把顶部导航首页链接改用比较简短的关键词做锚文字，不过用户体验不好，用户不习惯，不能肯定这个链接就是通往首页的。

要解决这个问题一个比较简单的办法，就是页面上第一次（或最前几次）出现的以“首页”为锚文字的链接，加上 NoFollow 属性禁止搜索引擎跟踪，然后在页面上其他适合放一两个关键词的地方，如页脚，以关键词为锚文字链接向首页。这样，页面上第一个搜索引擎可以跟踪的链接就变成了页脚上的链接，而且锚文字中包含目标关键词。

另外一种方法，就是使用 CSS 控制页面表现，使页面上看起来第一个出现的以“首页”为锚文字的首页链接，实际上在代码中并不是第一个出现。代码中第一次出现的是以关键词为锚文字的首页链接，但表现上是出现在页面底部。

4.14.9 深层链接

给网站深层页面，尤其是具体产品页面建设一些外部链接，不仅有助于使外部链接构成趋向自然，也有助于页面收录。不仅对被链接的产品页面收录有帮助，还对与之在链接关系上相邻的区域内的页面收录都有好处。

外部链接进入网站最多的是首页，搜索引擎蜘蛛跟随外部链接进入网站后，爬行和抓取的路线就像扔一颗石子到水中形成的波纹一样，从中心进入点向外扩散。从首页进入的蜘蛛扩散后，就爬向分类页面、子分类页面，然后是具体产品页面。

跟随外部链接从某一个产品页面进入的蜘蛛，同样有这样的扩散爬行路线。蜘蛛进入后，会向与之相连的前后页面、上级分类页面、相关产品页面等扩散。因此，给一些距离首页较远、不太容易被蜘蛛爬到的页面适当建设几个外部链接，可以有效地解决一个区块的所有页面收录问题。

4.14.10 分类隔离

仔细检查本章前面讨论的典型树形链接结构，不知读者是否能看出一些问题。这种典型树形结构对大部分网站来说是最优化的，但有的时候由于域名权重比较低，就算网站比较扁平，最终产品页面还是权重过低，无法达到搜索引擎蜘蛛爬行收录的最低标准。这时可以考虑彻底改变树形结构。

仔细观察标准树形结构可以看到一个潜在弱点：分类页面得到太多链接和权重。不仅首页直接链接到分类页面，分类页面之间互相链接，网站上所有最终页面也通过主导航系统链接到所有一级分类，以及一部分二级分类页面。也就是说，在权重分配上，级别高的分类页面和首页几乎差不多，得到了网站所有页面的链接及传递的权重。

对大部分网站来说，分类页面收录不成问题。分类页面积累的权重过高，反而使得最终产品页面获得的权重比较低。站长可以考虑把树形结构改为将不同分类进行分隔的链接结构。

在这种结构下，一级分类只链接到自己下级分类，不链接到其他一级分类。二级分类页面只链接回自己的上级分类，而不再链接到其他一级分类（包括其他一级分类下的二级分类）。同样，最终产品页面只链接回自己的上级分类页面，不再链接到其他分类页面。这样，分类之间形成隔离，首页权重将会最大限度地“灌入”到最终产品页面，而不是浪费在分类页面上。

这里所说的不链接到其他分类页面，既可以是真的取消链接，也可以通过禁止蜘蛛爬行的 JavaScript，Flash 等方式实现。

一些网站的实验证明，恰当使用这种方式可以使原本没有被收录的整个分类整体权重提升，达到被收录的最低标准。

要注意的是，这种方法只考虑收录，而没有考虑分类页面的排名问题。分类页面获得权重降低，也意味着排名力降低。

另外，更为重要的是，这种结构非常复杂，程序人员在处理哪些页面可以链接向哪些页面时必须非常小心，一不留神就可能使整个网站链接关系混乱。

这是比较难以掌握的方法之一，不到万不得已一般不要尝试。

4.15 CMS 系统

现在的网站绝大多数是通过 CMS 系统制作和管理的，SEO 人员优化网站就会受限于 CMS 系统本身的灵活性和功能。市面上有很多开源免费的 CMS 系统，如中文的 dede（织梦）、帝国，英文的 Joomla、Drupal 等。大部分 CMS 系统已经注意到搜索引擎友好问题。有的系统自身已经做了 SEO 方面的调整，有的需要安装插件。所以大部分现成的 CMS 系统具备一定的 SEO 基础。

不过，完美实现 SEO 效果的 CMS 系统，我还没有见到，经常需要修改代码才能完成一些功能。当然，自己开发 CMS 是最灵活的。下面讨论从 SEO 角度出发 CMS 系统应该具备哪些基本功能，SEOer 在选择或开发 CMS 时可以参考。

1．页面标题客制化

首先，CMS 系统应该自动生成比较优化的页面标题，如分类页面采用：

分类名称 – 网站名称

产品页面使用：

产品名称 – 分类名称 – 网站名称

或

产品名称 – 网站名称

CMS 系统还应该提供所有页面人工修改标题的客制化功能，在需要时站长可以根据关键词研究结果，重新撰写更恰当的标题。这是非常必要的功能。

2．URL 静态化及客制化

URL 静态化功能是必需的。通常 CMS 系统后台会有一个 URL 静态化选项，站长打开静态化功能，整个系统生成的页面就都使用静态化过的网址。静态化功能可能需要服务器支持某些模块，如 mod_rewrite。

好的 CMS 系统还应该提供 URL 客制化功能，站长可以自己定义所有的目录名及文件名。目录名通常在创建分类时就给出选项，站长可以使用自动生成的分类编号做目录名，也可以自己输入英文单词、汉语拼音等做目录名。创建产品或文章页面时，系统既可以自动按编号生成文件名，也允许站长输入最想要的文件名。

3．meta 标签客制化

这里所说的 meta 标签指关键词标签（keyword tag）和说明标签（description tag）。系统应该允许站长完全删除这两个标签，因为对搜索引擎排名几乎没有作用，如果不能生成恰当的标签，还不如不写。如果站长选择使用说明标签和关键词标签，应该可以选择生成格式或完全人工撰写。

4．正文撰写

现在的 CMS 系统创建产品或文章页面时都提供“所见即所得”编辑功能，用户可以在正文中加黑体、加图片、填写图片 ALT 文字、添加链接、选择链接锚文字等。

5．分类页面说明文字

除了首页应该给站长充分的灵活性、人工撰写说明文字外，分类页面也应该留出一个区域，让站长人工撰写最恰当的说明文字。这样才能避免分类页面上只有产品列表，没有独特相关的文字内容。

6．标签功能

站长发布产品或文章时可以输入标签，系统自动按标签聚合内容并生成标签页面。

7．消除复制内容

系统应该自动禁止抓取复制内容，如打印版本、各种不同排序（按价格、按时间、按热门程度等）页面。有的系统，如博客，天生有多种存档，除了主页面，还有按分类存档、按时间存档、按作者存档等。这些复制内容有利于用户浏览，却会给搜索引擎造成大量不友好的复制内容。CMS 设计者应该考虑到禁止抓取这些复制版本的功能，不然系统使用者很难在使用过程中做出改进。

8．301 转向

系统应该提供从任何一个页面做 301 转向到另外一个地址的功能，站长可以根据需要自由安排转向。没有这个功能，网站栏目的更改、删除就会很有局限性。

9．网站地图生成

好的 CMS 系统可以自动生成 XML 版本网站地图，站长只要向搜索引擎提交网站地图地址就可以了。

10．正确生成 H 标签

几乎所有 CMS 系统都会生成 H1、H2 标签，但正确生成 H 标签的并不多见。比如很多 CMS 系统在所有页面都把网站名称放在 H1 标签中，这实际上是不正确的。首页应该使用网站名称，或者允许客制化 H1 标签中的内容。分类页面应该把分类名称放入 H1 标签，而不是网站名称。产品页面，则应该把产品名称放入 H1 标签。

上面提到的只是搜索引擎友好的最基本要求。实际上还有很多细节，比如翻页的处理、相关文章的生成，很少有现成 CMS 系统能完美解决，往往需要自己开发软件系统，才能达到要求。

4.16　404 页面

用户访问网站上不存在的页面时，服务器通常应该返回 404 错误。如果站长没有在服务器端设置客制化的 404 页面，用户浏览器显示的将会是一个默认错误页面，如图 4-12 所示。

图 4-12　默认 404 错误页面

这样的错误页面用户体验肯定不会好。所有的主机都提供客制化 404 页面的功能，站长应该充分利用。

不能假设页面既然不存在，搜索引擎就不会来爬行这种网址。由于种种原因，网上任何地方都可能出现指向你的域名却写错目录或文件名的 URL，搜索引擎蜘蛛会跟踪这种错误 URL，访问不存在的页面。

4.16.1 404 错误代码

首先要注意的是，当页面不存在时，一定要返回 404 代码。有的服务器设置有问题，或者站长有意在页面不存在时还返回 200 状态码，也就是表示页面数据正常，这样将使搜索引擎认为网站上有大量重复内容，多个 URL 正常返回页面，但内容全是一样的。

页面不存在时，也不要返回任何转向代码。有的站长觉得既然页面不存在，就将用户 301 转向到首页，这也是搜索引擎不友好的设置，会让搜索引擎认为网站上有大量与首页内容相同的页面。

也不要使用 JavaScript 转向或 Meta Refresh 转向，尤其是时间比较短的如 10 秒以下的 Meta Refresh 转向。这些转向都使搜索引擎误以为页面存在，却返回重复内容。

不存在页面一定要确保正确返回 404 状态码。可以用后面 SEO 工具部分介绍的服务器头信息检测工具，任意输入一个肯定不存在的 URL，看服务器返回什么头信息，如图 4-13 所示。

```
#1 Server Response: http://www.chinamyhosting.com/nopage.html
HTTP Status Code: HTTP/1.1 404 Not Found
Date: Fri, 09 Apr 2010 18:31:12 GMT
Server: Apache/2.2.15 (Unix) mod_ssl/2.2.15 OpenSSL/0.9.8e-fips
Content-Length: 563
Content-Type: text/html; charset=iso-8859-1
```

图 4-13　确认服务器返回 404 代码

4.16.2 404 页面设计

客制化 404 页面设计上需要注意几点。

首先 404 页面要保持网站统一模板、设计风格、logo 及名称，不要让用户弄不清自己到了哪个网站上。

404 页面应该在最醒目的位置显示错误信息，明确提示用户，要访问的页面不存在。还可以加上几点可能性，如页面已删除、用户输入了错误的地址、链接中的地址错误、页面已转移到新的地址等。

错误信息下面还可以为用户提供几种点击选项，如网站地图，包括通往首页和重要分类页面的链接，建议用户可以访问的页面，还可以加上站内搜索框。

4.16.3 404 错误与外链

在这里提一个与 404 有关的非常简单有效的外部链接技巧。

站长可以在 Google 网管工具看到 Google 爬行了哪些不存在的网址，如图 4-14 所示。

通常 Google 蜘蛛不会没有原因地爬行这些不存在的页面，一定是有链接连向这些地址，而且大部分情况下是其他网站链接过来，却由于种种原因连向了错误地址，如图 4-15 所示。站长可以用 Google 网管工具检查有哪些外部网站连向了哪些不存在的页面。

图 4-14　Google 网管工具列出的 404 错误

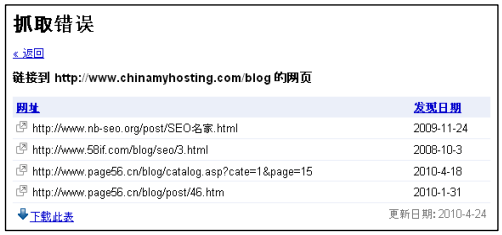


图 4-15　指向不存在页面的外部链接

这些外部链接连向不存在页面是个浪费，站长完全可以充分利用，或者联系对方站长改正错误地址，或者将错误地址做 301 转向到正确页面，或者干脆在本来不存在的 URL 上建一个页面，接收这些外部链接的权重。

第 5 章

页面优化

本章讨论页面本身可以优化的元素。页面和网站结构是 SEO 自己可以控制的，优化好这两方面就给网站 SEO 打下了良好的基础。

5.1 页面标题

页面标题是包含在 Title 标签中的文字，是页面优化最重要的因素。

标题标签的 HTML 代码格式如下：

```
<head>
<title>SEO 每天一贴 - Zac 的 SEO 博客，正在写书，暂时每月两三贴</title>
…
</head>
```

用户访问时，页面标题文字显示在浏览器窗口最上方，如图 5-1 所示。

在搜索结果页面上，页面标题是结果列表中第一行文字，是用户浏览搜索结果时最先看到的、最醒目的内容，如图 5-2 所示。

图 5-1　浏览器窗口显示的页面标题

图 5-2　搜索结果列表中的标题

建议 Title 紧接着写在<head>之后，然后再写其他标签和代码，尤其不要在中间插上一大段 JavaScript 程序，这样搜索引擎可以迅速找到标题标签。

页面标题的优化要注意下面几点。

5.1.1 独特不重复

即使在同一个网站内，主题相同，不同页面具体内容不会相同，页面标题也不能重复，每个页面都需要有自己独特的标题标签。不同页面使用相同的标题是不了解 SEO 的站长经常犯的错误之一，就连一些重要的 IT 人物和大网站也经常会犯这个错误。

最常见的重复标题是忘记写标题标签，使用了编辑软件新建文件的默认标题。在中文页面上经常表现为“未命名文件”，在英文里则显示为“Untitled Document”。如图 5-3 所示，在 Google 搜索 Title 中包含“Untitled Document”的文件，返回 8500 多万个结果。

图 5-3　大量以“Untitled Document”为标题的页面

搜索串中加了一个“的”字，目的是只显示中文页面。通过图 5-3 可以看到，很多教育机构的页面经常不写标题标签。

再看如图 5-4 所示的比尔・盖茨的博客。

图 5-4　比尔・盖茨博客的重复标题

虽然写了标题标签，但标题都是一样的。很多企业网站也是如此，如图 5-5 所示。

标题是页面优化第一位的因素，是搜索引擎判断页面相关性最重要的提示。不同页面使用重复标题是极大的浪费，用户体验也不好，搜索用户无法从页面标题看出这个页面到底是什么内容。

小网站经常需要人工撰写最合适的标题标签。大型网站页面数量众多，不可能人工撰写，需要通过程序调用页面上特有的内容生成标题。最常见的是最内页直接调用产品名称或文章标题，分类页面使用分类名称，首页建议人工撰写。

图 5-5　企业网站的重复标题

使用重复标题经常是因为疏忽。要检查网站上是否有重复标题，可以使用 Google 站长工具，故障诊断-HTML 建议部分会列出使用了重复标题的页面，如图 5-6 所示。

图 5-6　Google 网管工具列出有重复标题的页面

可以看到，我的博客很多存档页面有重复标题。这里提醒读者一下，我的博客“SEO 每天一贴”并不是一个 SEO 范例，有很多该优化的地方没有优化，标题的撰写只是一例，读者如果仔细研究会发现更多。原因是我的博客没有商业性，排名和流量对这个博客没有什么意义。所以不要盲目模仿我的博客，SEO 方法请以本书中所介绍的为准。

有的时候要生成独特标题并不是一件简单的事，比如电子商务网站的分类页面，同一个分类下产品数量比较多时，产品列表页面必然需要翻页，大型网站可能会翻几十到几百页。这些分类页面的标题通常都是“分类名称 – 网站名称”格式。这时程序员就需要在标题中加入页数，使翻页页面标题标签不同。分类第一页不必加页号，从第二页开始页面标题最前面加上“第二页”、“第三页”等文字，形成的标题大致如下：

第一页：

分类名称 – 网站名称

第二页：

第二页 – 分类名称 – 网站名称

5.1.2 准确相关

这一点不言自明。每个页面的标题都应该准确描述页面内容，用户看一眼就能知道将访问的页面大致是讲什么的，搜索引擎也能迅速判断页面的相关性。

很自然，准确描述页面内容，就必然会在标题中包含目标关键词。唯一要注意的是，不要在标题中加上搜索次数高、但与本页面无关的关键词。这种方法七八年前还管用，现在早已没有任何效果，弄不好还会被认为是作弊。

5.1.3 字数限制

从纯技术角度说 Title 标签可以写任意长度的文字，但搜索结果列表页面标题部分能显示的字数有一定限制。百度最多显示 30 个中文字符，Google 显示 65 个英文字符，转成中文大概是 32 个字符。Title 标签中超过这个字数限制的部分将无法显示，通常在搜索列表标题结尾处以省略号代替。

前面介绍过，用户搜索的关键词会在显示的标题中被加红色高亮。如果搜索词位于 Title 标签后面超出显示字数部分，百度会显示 Title 中后面包含关键词那部分文字，前面加省略号。如搜索“site:dunsh.org 论坛”，有的页面显示如图 5-7 所示。

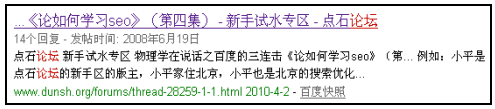


图 5-7 百度有时候因搜索词位置显示标题后半部分

页面的实际标题如图 5-8 所示。

图 5-8 页面标题长度超过百度能显示的字符数

所以标题标签不要超过 30 个中文字，为保险起见最好不要超过 25 个中文字。实际上为了提高用户体验及突出目标关键词，我通常建议标题标签以 15～20 个中文字比较适合。

标题标签之所以不宜超过搜索结果列表所能显示的字数限制，主要有三点原因：

（1）搜索结果列表标题不能完整显示时，结尾会被切断，而站长又不容易预见在什么地方被切断。如图 5-9 所示，结尾处的字被拦腰截断，用户不能轻松知道应该是哪个词。

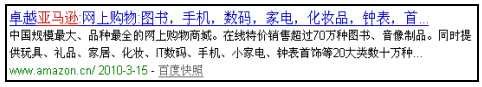


图 5-9　标题过长被截断

（2）虽然超过显示字数限制的标题标签并不会引来搜索引擎的惩罚，但超过显示字数的部分搜索引擎会降低权重，对排名作用很小。

（3）标题越长，在不堆积的前提下，无关文字必然越多，不利于突出关键词相关性。

5.1.4　简练通顺，不要堆砌

堆积关键词也是初学 SEO 的人很容易犯的错误，为了提高相关性，在标题中不自然地多次出现关键词。比如把标题写成：

女装|女装批发|女装零售|女装批发零售|××女装网

其实这样的页面只要写成：

女装批发零售|××女装网

就可以了。

SEO 人员不仅要考虑搜索引擎，还要考虑到用户，不能让用户一看标题就感觉这是一个为关键词而关键词的页面，而需要把标题写成一个正常通顺的句子。在这一点上百度百科是最好的教材，如图 5-10 所示。

图 5-10　百度百科标题简洁

标题干净利索，一句废话没有。SEO 人员都知道现在百度百科在百度本身及 Google 排名都非常好，简练通顺的标题很可能是重要原因之一。

5.1.5 关键词出现在最前面

在可能的情况下，目标关键词应该出现在标题标签的最前面。经验和一些统计都表明，关键词在标题中出现的位置与排名有比较大的相关性，位置越靠前，通常排名就越好。

对网站内页来说，通常是按：

产品名称/文章标题–子分类名称–分类名称–网站名称

的格式生成标题标签。

标题标签可以理解为倒置的面包屑导航，比如页面所在位置的面包屑导航是：

首页>电子产品>数码相机>索尼数码相机

那么页面的标题就写成：

索尼数码相机 – 数码相机 – 电子产品 – ××电器网

这个页面的目标关键词是索尼数码相机，出现在标题最前面。考虑到字数限制和突出目标关键词，也可以拿掉分类名称，把标题缩减为：

索尼数码相机 – ××电器网

5.1.6 吸引点击

提高关键词相关度和排名是带来搜索流量的一方面，点击率是搜索流量的另一个决定因素。标题标签能够吸引用户目光，让用户欲罢不能，非要点击看个究竟，才能达到最好的 SEO 效果。比如这样的标题属于中规中矩的：

减肥茶 – 减肥产品 – ××减肥网

如果写成这样点击率就可能有所上升：

减肥茶：无须节食，无须运动，快乐减肥 – ××减肥网

减肥茶 10 天减轻体重 5.5 公斤，真实用户证言 – ××减肥网

减肥茶免费样品大赠送 – ××减肥网

索然无味的页面标题即使出现在搜索结果前面也可能很少有人点击，好的标题即使排在后面也可能点击率更高，带来的流量也更大。单纯的排名没有意义，有人点击才有效果。

好的 SEO 除了要了解基本优化技术，也得研究用户心理、文案写作等。

5.1.7 组合两三个关键词

一般来说，一个页面最多针对三四个关键词进行优化，不宜再多。

前面抓图显示的亚马逊书店把“图书，手机，数码，家电，化妆品”等都放在首页 Title 中，如果亚马逊网站这样做真的是想把这么多关键词定位在首页，那么这不是一个好的策略。普通企业和站长不能看到品牌网站这么做就跟着学。亚马逊这种级别的网站流量肯定不能靠 SEO，页面优化得怎样并不那么重要，而且大部分东西不是从 SEO 角度出发的。

要针对四五个或更多关键词优化页面，往往使页面内容不能突出，文案写作上也会顾虑太多，无法自然融入这么多关键词。一个页面如果需要优化很多关键词，往往是关键词研究没有做好，没有正确把关键词分组并分布到各个栏目和内页中。

有的时候，两三个关键词只要并排放在标题标签中即可。比如你可能在做关键词研究时发现“韩版女装”和“韩国女服”这两个词搜索次数都不少（只是假设），而这两个词的意义显然完全一样，在标题中写成这样就可以了：

韩版女装，韩国女服 – ××服装网

有的时候三四个关键词还可以组合为更通顺的一个词组。在写这一节时，刚好浏览 zaccode.com 网站看到一个朋友问，如果页面想优化“SEO 技术”，“SEO 教程”，“SEO”，“免费 SEO”这四个词，该怎么写标题？我的回答是：

免费 SEO 技术教程

这是一句很通顺的话，而且搜索引擎可以拆分、组合出所有目标关键词。后面的案例部分还有更多实际的例子。

当然，这样组合出的关键词有时不是完整匹配，比如上面例子中的“SEO 教程”就不是完整匹配出现在标题中的。但与用户体验相比，为了刻意完整匹配而用一个不自然的标题，如这样：

SEO-SEO 技术 – SEO 教程 – 免费 SEO

我建议还是首先考虑用户体验。而且不自然的文字可能被搜索引擎认为有作弊嫌疑。

把多个关键词组合为一句更为通顺的词组或句子，没有适用于所有情况的规则，只能自己发挥创意，不断练习，做久了自然熟能生巧，既能包含主要关键词，又让用户觉得没有优化痕迹。

5.1.8 公司或品牌名称

通常把公司或品牌名称放在标题的最后是个不错的做法。虽然用户在搜索公司名称、网站名称时，一般只有首页才会被排在最前面，分类页、具体产品页等内页既没有必要，也没有可能针对品牌名称做优化，但是品牌名称如果多次出现在用户眼前，就算没有点击也能让用户留下更深的影响。

有的时候品牌名称比任何关键词都重要，写标题时关键词就不是考虑的重点了。例如搜索“网上银行”的次数不少，但银行网站首页标题如果写成这样：

网上银行，网上银行查询转账 – 招商银行

那就真是捡了芝麻，丢了西瓜。

5.1.9 连词符使用

标题标签中词组之间需要分隔时，既可以使用空格，也可以使用连词符 - | >· _ 等。这些连词符在排名上没有什么区别，用哪个都一样，但是在显示出来的视觉效果上可能有不少差别。选用哪个连词符，只取决于你认为哪种符号能让用户看得最清楚、看得最舒服。个人认为 ->| 都是不错的选择。

5.1.10 不要用没有意义的句子

标题标签是最宝贵的优化资源，不要浪费在没有意义的句子上。诸如“公司欢迎您”，“用户是我们的上帝”之类，既不能提高相关性，也不能吸引用户目光和点击量。

5.1.11 noodp 标签

网站被开放目录收录时，Google、雅虎、必应有时候会取开放目录中的标题作为搜索结果列表中的页面标题，而不使用页面本身的 Title 标签内容。而开放目录中的标题是由编辑确定的，不一定是站长最想要的标题。站长可以使用 noodp 标签强制搜索引擎不要使用开放目录标题。

noodp 标签放在 HTML 的 head 部分，代码如下：

```
<meta name="googlebot" content="NOODP">
```

标题标签是 SEO 过程的重点之一，建议大家重要页面都人工撰写，如首页、分类或栏目页、热卖产品页面等。而撰写标题时又需要做最起码的关键词研究，才能知道哪几个关键词搜索次数最多或性价比最高。大型网站很可能有数百个分类及子分类，人工撰写标题是一件费时、费力、但一定有效果的工作。

5.2 描述标签

描述标签是 HTML 代码中 Head 部分除标题标签外与 SEO 有关系的另一个标签，用于说明页面的主体内容。描述标签的代码格式为：

```
<head>
……
<meta name="description" content="SEO 每天一贴研究搜索引擎优化 SEO 技术，网络
营销及电子商务思考。正在写一本 SEO 方面的书，所以不能每天一贴了，会尽快恢复。" />
……
</head>
```

描述标签的重要性比标题标签低很多。描述标签中的文字并不显示在页面可见内容中，用户只有查看源文件和在搜索结果列表中才能看到描述标签里的文字。

现在主流搜索引擎排名算法都已经不使用描述标签，所以描述标签对关键词排名没有影响，但是对点击率有一定影响，因为大部分情况下，搜索结果列表中的页面摘要说明就来自描述标签。

除了描述标签外，搜索结果列表中的页面说明还可能来自另外两个来源。一是搜索引擎自动抓取页面可见文字中的相关段落。尤其描述标签中不包含所搜索的关键词时，搜索引擎经常从页面可见内容中动态抓取包含搜索词的部分显示为说明文字。

另外一个来源是重要网站目录。网站如果被雅虎或开放目录等收录，搜索引擎也可能会使用雅虎或开放目录的说明文字作为页面说明。与标题标签一样，站长可以使用 noodp 标签禁止搜索引擎使用开放目录的说明文字。在一些特殊情况下，搜索引擎也可能组合使用上面三种说明文字来源。

说明文字中的搜索词也以高亮显示，以便于提高点击率。

当搜索引擎动态抓取可见内容作为说明文字时，通常会从不同段落提取文字，因而句子之间会出现省略号，而且在说明文字的前后、中间都可能出现省略号，如图 5-11 所示。

到底SEO是什么？ – 搜索引擎优化SEO每天一贴
2009年5月15日 ... 做几年SEO之后，仔细回想每天进行的工作也会经常迷惑，到底SEO是什么？
... 所以看一圈下来，你会发现不知道到底SEO是什么了。我们所做的大部分是常识 ...
www.chinamyhosting.com/seoblog/.../whats-seo-really/ - 网页快照 - 类似结果

图 5-11　动态提取的说明文字

搜索引擎动态提取说明文字时，站长就无法控制，所显示的说明文字有可能在重要的地方被截断。所以虽然描述标签不用于排名计算，在可能的情况下还是建议写上描述标签，以便控制页面的说明文字。

在下面几种情况下，搜索引擎更可能动态抓取它认为合适的说明文字：

- 描述标签包含大量堆积关键词。
- 描述标签与标题标签内容重复。
- 描述标签只是关键词的罗列，不能形成通顺的句子。
- 描述标签不包含用户所搜索的关键词。

在描述标签的写作上，大部分标题标签写作要点依然适用，比如文字需要准确相关、简练通顺、吸引点击，不要堆积关键词，每个页面有自己独特的描述标签，包含目标关键词等。

与标题标签不同的是，搜索结果列表说明文字部分通常会显示 77 个中文字符，Google 英文结果显示 156 个英文字符，比标题标签写作空间大一点。为保险起见，建议描述标签不要超过 70 个中文字。

小型网站站长可以人工撰写描述标签，用一两句通顺的句子说明页面主题。大中型网站则不可能人工撰写，通常可以采取两种方法自动产生。一是从页面正文中提取一部分，一般会提取第一段文字中的内容。二是从产品重要信息生成，如产品名称、品牌、型号、价格、颜色、生产商等。在可能的情况下，大中型网站的首页及主要分类页面也可以考虑人工撰写描述标签。

最后要说明的是，如果不能生成恰当通顺、不重复的描述标签，那么就不要写描述标签。

5.3 关键词标签

关键词标签是 HTML 代码 Head 部分看似与 SEO 有关、但目前实际上对 SEO 没有任何影响的标签。关键词标签的代码格式是：

```
<head>
……
<meta name="keywords" content="关键词 1, 关键词 2, 关键词 3……" />
……
</head>
```

关键词标签本意是用来指明页面的主题关键词。十几年前，在关键词标签中重复关键词，甚至加上无关的热门关键词曾经是排名的重要方法。也正因如此，站长们发现了这个方法，立即引来了大规模滥用。所以现在的主流搜索引擎无一例外都没有在排名算法中考虑关键词标签的内容。

几年前关键词标签还可以用来放异体字、错拼单词等，但是现在的搜索引擎都有错字错拼提示功能，关键词标签的这个用法也失去了意义。

也许以后搜索引擎还会把关键词标签计入算法，但目前的建议是，不用浪费时间写关键词标签。

HTML 代码 Head 部分的其他 meta 标签都对 SEO 没有影响。

5.4 正文中的关键词

页面正文中出现关键词是必要的，也是必然的。

5.4.1 词频和密度

正文中的关键词牵扯到两个概念。一个是词频，也就是关键词出现的次数。另一个是关键词密度，也就是关键词出现次数除以页面可见文字总词数，或者说关键词密度是规范化后（考虑正文长度后）的词频。

最初的搜索引擎算法认为，关键词出现次数越多，也就是词频越高，页面与这个关键词越相关。但词频概念没有考虑内容长度。页面正文如果是 1000 个词，显然关键词词频很容易比 100 个词的页面高，但并不必然比 100 个词的页面更相关。用关键词出现次数除以总词数，得到关键词密度，是更合理的相关性判断标准。

现在的搜索引擎算法已经比简单词频或密度计算复杂得多。站长可以很容易地人为提高词频和密度，页面价值却不一定更高。所以页面对关键词的相关性与词频和密度已经没有直接联系，SEO 人员不必太在意。

一般来说，篇幅不大的页面出现两三次关键词就可以了，篇幅比较长的页面出现 4～6 次也已经足够，千万不要堆积关键词。

初学 SEO 的人常常问关键词密度是多少最合适。其实关键词密度在现在的搜索引擎算法中已经是一个不太相关的概念，SEO 人员完全不必关注。观察排名在前面的页面，我们会发现其中既有密度低到 1%或 2%的页面，也有高到 20%的页面。只要自然写作，页面中必然出现几次关键词，这就已经完成优化了。

5.4.2 前 50～100 个词

正文前 50～100 个词中出现的关键词有比较高的权重，通常建议第一段文字的第一句话就出现关键词。

实际上这也就是自然写作的必然结果。和写议论文一样，页面的写作也可以分为论

点、论据及最后的总结点题。文章最开头首先需要点明论点，也就自然地包含关键词。接下来的论据部分再出现两三次，结尾点题再次出现关键词，一个页面的可见文字优化就完成了。

5.4.3 关键词变化形式

写作页面内容时可以适当融入关键词的变化形式，包括同义词、近义词、同一件事物的不同称呼等。比如电脑和计算机是同义词，可以在页面中交叉出现。笔记本电脑和膝上型电脑也是同义词，减肥和瘦身是近义词，主机、服务器、网站空间基本上说的是一回事，这些词都可以在主体内容中交叉出现。

英文页面优化中还有一种特殊关键词变化形式，就是同一个词根所生成的各种形式的单词，如 work 是词根，它的变化形式包括 works，worker，workers，working，worked 等。这些变化形式交叉使用，使搜索引擎能更快速准确地提取页面主题，并且这些变化形式都有可能被用户所搜索，组合出更多搜索词。

5.4.4 关键词组临近度

标题标签和正文写作时应该注意目标关键词组的临近度，也就是说，关键词可以被分词时，在页面上应该完整、按顺序出现关键词组几次，尤其是重要位置，如本节提到的第一段文字，其他节说明的 H 标签、黑体、ALT 文字等。如目标关键词是“SEO 方法”，就要在页面上完整出现“SEO 方法”，而不是把“SEO”和“方法”分开，分别出现在页面上。

百度对关键词完整匹配出现的要求比 Google 高些。

5.4.5 词组的拆分出现

搜索词可以被分词时，不仅搜索词要完整匹配出现在页面最有权重的位置，被拆分后的词还可以各自单独出现在正文中几次。

假设目标关键词是“点石互动论坛”，以百度为例，这个词会被分词为“点石互动”和“论坛”两个词。页面重要位置不仅要完整匹配出现“点石互动论坛”六个字，建议“点石互动”和“论坛”也可以分别单独（不连在一起）出现几次。

5.4.6 语义分析

算法和人很不一样的地方是，人可以直接理解词的意思、文章的意思，算法不能理解。人看到“苹果”这两个字就知道指的是那个圆圆的、有汁的挺好吃的水果，搜索引擎却不能从感性上理解什么是苹果。

但搜索引擎可以掌握词之间的关系，这就涉及语义分析。SEO 业界很热烈地谈论过一阵潜在语义索引（Latent Semantic Indexing）。潜在语义索引研究的是怎样通过海量文献找出词汇之间的关系，当两个词或一组词大量出现在相同文档中时，这些词之间就被认为是语义相关的。

举个例子，电脑和计算机这两个词在人们写文章时经常混用，这两个词在大量的文档、网页中同时出现，搜索引擎就会认为这两个词是极为语义相关的。实际上它们是同义词。

潜在语义索引并不依赖于语言，所以“SEO”和“搜索引擎优化”虽然一个是英语，一个是中文，但这两个词大量出现在相同的网页中，虽然搜索引擎还不能知道搜索引擎优化或 SEO 指的是什么，但是却可以从语义上把“SEO”，“搜索引擎优化”，“search engine optimization”，“SEM”等词紧紧地连在一起。

再比如“苹果”和“橘子”这两个词，也经常出现在相同的文档中，虽然紧密度低于同义词，但搜索引擎可以判断这两个词是有某种语义关系的。

搜索引擎有没有使用潜在语义索引，至今没有定论，没有搜索引擎承认或否认。但一个事实是，2002 年 Google 买下了拥有潜在语义索引专利的一家公司 Allied Semantic。

这种语义分析技术可以在搜索引擎优化上给我们一些提示。

从大概 2004 年开始，Google 排名有一个现象，搜索某个关键词，排在前面的网页有时甚至并不含有所搜索的关键词，除了链接锚文字外，很有可能是潜在语义索引在起作用。比如搜索“电脑”，排在前面的有可能包括一篇只出现“计算机”却没出现“电脑”这个词的页面，因为搜索引擎通过语义分析知道这两个词是紧密相关的。

在进行网页写作的时候，不要局限于目标关键词，应该包含与主关键词语义相关、相近的词汇，以支持主关键词。

假设优化“培训”这个词，要加强相关性，并不是多出现几次“培训”这么简单，还可以通过出现其他语义相关词汇。和“培训”语义相关的词包括：老师，学生，理论，技巧，师父，徒弟，考试，实践，学校，师资，教育等。由于语义分析的因素，这些词在搜索引擎眼里是比较相关的，在写内容时，就要把这些词符合逻辑、符合语法地融入进去。

5.4.7 分类页面说明文字

大部分网站首页及最终产品或信息页面优化并不困难，有足够的内容可以安排关键词，但分类或频道页面往往被忽视。分类和频道页面最常见的内容就是产品或文章列表，而产品名称或文章标题实际上都是产品或具体信息页面的重复内容，这就造成分类页面缺少自己的独特内容。

而分类或频道页面的目标关键词往往是排在第二层次的搜索次数比较多的词。要想充分优化这类关键词，就需要人工撰写分类或频道页面的说明文字，最好有至少两三段说明文字。

从 SEO 角度看正文中的关键词优化，最主要的还是自然写作。一些单页面网站，站长可以精雕细琢关键词分布、用词，但绝大部分网站制作时，无论是编辑还是专业 SEO 人员，都不可能时刻想着关键词密度、同义词、语义相关词、词组的拆分等细节，那样做效率太低。其实只要自然写作，上面讨论的这些方面就自然而然地做到八九不离十了。

如果头脑中能有关键词在重要位置的分布、同义词近义词使用、语义相关、词组的拆分组合等观念，写作时从有意识到下意识地运用这些方法，再好不过了。

更重要的是针对用户及转化率的写作。对用户有帮助的原创内容，强调卖点，消除购买顾虑，排版格式清晰易读，适当的地方加上图片美化等，都可以提高转化率。

5.5 H 标签

H 标签相当于正文标题，是关键词优化的另一个页面元素，通常认为重要性仅次于页面 Title。近两年 Google 给予 H 标签的权重有所降低。

H标签按重要性分为六层，从 H1 到 H6。H1 的 HTML 代码是：

```
<h1>SEO 每天一贴</h1>
```

H1 最重要，H6 重要性最低。在 H1 和 H2 标签中融入关键词，有助于提高相关性。H3 以下的标签权重已经很低，和普通页面文字相差不多了。

H 标签对应于文章的正常结构。文章一定会有一个标题，应该使用 H1 标签，其中包含最重要的关键词。文章中出现小标题，使用 H2 标签，可以包含辅助关键词或相关词语。如果还有更小的标题，可以使用 H3。再低层的标题意义不大，不仅使文章结构过于复杂，搜索引擎给予的权重也很低了。

典型的 H 标签使用结构：

```
<h1>什么是 SEO？ </h1>
<p>第一段文字…</p>
…
<h2>SEO 是搜索引擎优化</h2>
<p>扩展、论述上面 h2 中内容</p>
…
<h2>SEO 是策略</h2>
<p>扩展、论述上面 h2 中内容</p>
…
<h2>SEO 不是作弊</h2>
```

```
<p>扩展、论述上面 h2 中内容</p>
…
<p>最后点题</p>
```

要注意的是，H 标签在视觉表现上常常是黑体，但其语法意义与黑体完全不同，不要在页面上滥用。一般来说一篇文章只有一个标题，H1 也只出现一次。大量使用 H1、H2，反倒使得关键词不能突出。整篇文章都放进 H 标签中，就和完全没用 H 标签一样了。

5.6 ALT 文字

ALT 文字是指图片的替换文字。代码如下：

```
<img src="images/default/logo.gif" alt="点石论坛" />
```

在某些情况下，比如用户浏览器禁止显示图片，或由于网络等原因图片文件没有被下载完成，以及视障人士使用的专用浏览器，导致图片不能被正常显示，图片 ALT 属性中的文字将被显示在页面上，如图 5-12 所示。

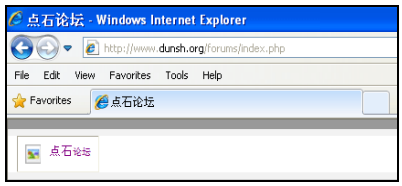


图 5-12　图片未显示时出现 ALT 文字

某些浏览器，如 IE7 以前的 Internet Explorer，鼠标放到图片上时，ALT 文字也会被显示出来。

图片 ALT 文字中出现关键词对页面相关性也有一定影响，近一两年 ALT 文字重要程度还有所提高，所以在图片 ALT 属性中以简要文字说明图片内容，同时包含关键词，也是页面优化的一部分。

与页面 Title 一样，ALT 文字中不要堆积关键词，只要出现一次关键词即可。

图片做成链接时，ALT 文字就相当于文字链接的锚文字。所以网站左上角出现的公司 logo，应该在 alt 文字中包含首页目标关键词。logo 一般都链接到首页，而且通常是页面上出现的第一个连到首页的链接，ALT 文字中包含关键词，就相当于文字链接锚文字中包含关键词。

图片文件名中包含关键词对图片本身出现在图片结果中，以及所在页面的相关度也有一点影响。这种优化主要用在英文网站上。中文网站用中文关键词做图片文件名，和关键词出现在 URL 中类似，作用不大，有的时候得不偿失。

5.7 精简代码

在搜索引擎工作原理一节中介绍过，搜索引擎预处理的第一步就是提取文字内容。SEO 人员应该尽量降低搜索引擎提取文字内容的难度，也就是精简 HTML 代码，使真正的文字内容比例提高，尽量减少 HTML 格式代码。从某种意义上来说，格式代码对关键词来说都是噪声，精简代码就是提高信噪比。

常见的可以精简代码的地方如下：

（1）使用 CSS 定义文字字体、颜色、尺寸及页面排版。有很多网站既使用 CSS，又在可见文字部分用 style 或 font 再定义一遍字体、尺寸等，这是完全没有必要的冗余代码。

（2）使用外部文件。将 CSS 和 JavaScript 放在外部文件中，页面 HTML 中只要放一行代码进行调用就可以了。如果我们查看一些网站源文件，经常可以看到大片大片的 CSS 及 JavaScript 代码，而且 JavaScript 代码还经常出现在 HTML 最前面，这就使真正有用的文字部分被推到后面。

（3）减少或删除注释。代码中的注释只是给程序员或页面设计人员的提示，对用户和搜索引擎来说毫无作用，只能成为噪声。

（4）减少表格，尤其是嵌套表格。现在的网页大多使用 CSS 排版，表格使用大大减少。但有的时候使用表格展现是必需的，也不必刻意完全避免，只要不出现多层嵌套表格，产生大量无用代码就可以了。

这里说明一下文件大小限制。Google 技术指南曾经建议，HTML 文件最好限制在 100KB 以下，页面上链接数在 100 个以下。百度目前建议 HTML 文件不要超过 128KB。其实现在的搜索引擎已经完全可以抓取大得多的文件，甚至一两兆的文件也没有问题。

不过在可能的情况下，还是应该尽量使文件越小越好。虽然搜索引擎可以抓取很大的文件，但可能不索引整个文件，而只索引文件前面一部分内容。通常，页面内容应该是集中统一的，索引前半部分就代表了文件内容主题。文件很大时，索引整个文件既没有必要，也是很大的资源浪费。文件过大，再加上大量冗余格式代码，可能使实质内容被推到实际被索引的部分之外。

5.8 内部链接及锚文字

网站架构一章讨论过，内部链接对爬行和收录有非常重要的意义。内部链接对页面关键词相关性也有影响，最主要的就是在内部链接中使用锚文字。

锚文字是告诉搜索引擎被链接页面主题内容的最重要依据之一。外部链接锚文字大部分是无法控制的，内部链接锚文字则完全由站长控制。锚文字中出现完全匹配的关键

词，有助于提高链接目标页面的相关度，以及发出链接页面的相关度。当然在这方面还要避免过度优化。除了一部分使用完全匹配关键词做锚文字外，最好有一部分锚文字具有自然多样性。

锚文字出现的位置不能集中在导航或页脚中，而需要分散在正文中。在页脚加上很多重要页面的链接，锚文字使用完全匹配关键词，曾经是一种很流行的优化方法，效果也曾经很明显。不过近一两年这种页脚过度优化常常是排名惩罚的原因之一。

5.9 导出链接及锚文字

链接对搜索引擎排名的重要性被越来越多的站长所了解和重视，造成很多网站惜链如金，不愿意导出链接到其他网站上。实际上导出链接到外部网站对发出链接的页面相关性也有一定影响。

比如说一个页面以“减肥方法”为锚文字链接到另外一个网站上的页面，一方面说明被链接的页面应该是在谈减肥方法，另一方面说明发出链接的页面本身也应该是与减肥方法相关的。不然一个谈笔记本电脑的页面，有什么原因要链接到一篇关于减肥方法的文章呢？所以链接到相关、高质量的外部网站，也有助于提高页面本身的相关性。

在搜索引擎链接原理部分我们讨论过枢纽和权威页面的关系，得到链接的往往是权威页面，而指向权威页面的往往是枢纽页面。如果不能成为权威页面，第二选项就是成为枢纽页面。导出外部链接就是成为枢纽的方法。

5.10 W3C 验证

页面 HTML 代码兼容性相差很大，各种操作系统的不同浏览器，甚至相同浏览器的不同版本解析 HTML 时都不同，越来越多的人在考虑 W3C 验证对 SEO 有什么影响。从经验和观察来看，W3C 验证通过与否对页面排名没有明显影响。只要页面没有严重错误，使搜索引擎无法提取文字内容，就不必太在意 W3C 验证。

实际上绝大部分页面都无法百分之百通过 W3C 验证，搜索引擎也非常清楚这一点。SEO 人员可以尽量更正验证报告中的错误，警告信息通常可以忽略。

5.11 黑体及斜体

黑体、斜体是页面文字很早就使用的格式，将关键词设为黑体或斜体有一点点强调作用，搜索引擎也给予黑体、斜体中的文字比普通文字多一点权重，不过权重并不大。尤其是斜体使用需要很慎重，中文斜体对用户不友好，不容易辨认。在可能的情况下，适当使用黑体有些作用，但这属于非常细节的地方，不必过于关注。

黑体有时有助于帮助分词。比如为避免搜索引擎把“搜索引擎优化”这几个字分词为“搜索”、“引擎”、“优化”三个词，可以把“搜索引擎优化”全部设为黑体，帮助搜索引擎理解，这六个字实际上是不应该分开的一个词。

5.12 页面更新

对某些有时效性的网站来说，如博客和新闻网站等，页面更新也常能提高排名。Google 和百度都有这样的现象，刚发布的文章很快能有比较好的排名，但过几天后排名会下降。更新内容使搜索引擎认为内容有时效性，更相关。

页面更新频率也是吸引搜索引擎蜘蛛返回抓取的因素之一。

5.13 Google 沙盒效应

所谓沙盒效应（Sandbox Effect），指的是新网站在 Google 很难得到好的排名，无论怎么优化这个网站。换句话说，一个新的网站，可以有很丰富的、相关的内容，可以有大量外部链接，网站既搜索引擎友好，也用户友好，所有一切都优化得很好，但是在一段时间之内，就是很难在 Google 得到好的排名，尤其是竞争比较大的关键词。

这个现象最早是在 2004 年 3 月开始被注意到的。Google 的 Sandbox 有点像给予新网站的试用期，在这段试用期内，新网站几乎无法在搜索竞争比较激烈的关键词时得到好的排名。

沙盒效应更多地发生在竞争比较激烈的关键词上。那些不太商业的，竞争比较小的关键词，发生沙盒现象的机会就比较小。进入沙盒的网站，搜索竞争大的主关键词找不到，搜索竞争比较小的关键词时则排名正常，收录也正常。

有的新网站一开始就有非常强的外部链接，如某些成为社会热点的网站，也可能不进入沙盒。

有时候老网站也可能因为短时间内增加大量外部链接而进入沙盒。

沙盒本身不是一个独立的、把所有新网站排名暂时调后的过滤算法，而是很多其他排名因素所造成的一种效应或现象。比如，Google 把域名年龄、链接的年龄、链接页的历史情况、网站获得链接的速度都考虑在排名算法内，这些与时间有关的因素组合起来，就可能对新网站产生沙盒现象。

大部分人认为 Google 之所以会考虑产生 Sandbox 效应的时间因素，是为了清除垃圾网站。通常垃圾网站建站后会快速买大量链接，得到好的排名，赚一笔钱后，作弊手段被发现了，网站被删除或被惩罚，但是作弊者也不在乎，这个域名也就被放弃了，转而开始做另外一个新的网站。Google 有了 Sandbox 以后，一部分垃圾网站制造者没这么多

耐心等他的网站从沙盒里出来，所以也就不建垃圾网站了。

通常沙盒效应会维持几个月，长的也有一两年的。关键词竞争不大的网站在沙盒里时间会短一些。行业竞争越高，沙盒效应越长。

虽然沙盒现象首先在 Google 排名中被观察到，也通常被称为 Google 沙盒，其实其他搜索引擎也有类似现象，只不过没有 Google 这么明显。

网站进入沙盒，SEO 人员该怎么办呢？首先要放宽心，因为从根本上说，站长没有好办法逃避沙盒现象。随着时间的推移，过大概 6 个月以后，新网站或新域名自然会从沙盒里出来。

同时，当网站在 Sandbox 沙盒里时，应该利用这段时间增加网站内容，建设外部链接，踏踏实实做好网站。实际上，沙盒效应对很多网站运营者来说可能是一件好事。因为在大概半年的时间里，只能把精力放在网站内容上面，而不会去考虑排名。从长远来看，一旦走出沙盒，外部链接的年龄足够长了，被记入算法当中，网站也有了足够的内容，排名会有一个质的飞跃。

大家可以看到，本书中“页面优化”这一章篇幅是最简短的，因为页面上能优化的地方就这么多，优化方法也相对固定。只要在重要的地方，如 Title、H1、alt 文字、链接锚文字等地方出现关键词，避免在任何地方堆积关键词就可以了。相比外部链接及网站结构来说，页面优化是比较简单的部分。在搜索引擎排名算法中，页面本身的优化只占 20%～30%的比重。

第 6 章

外部链接建设

网站优化分为网站内优化及网站外优化两部分，网站内优化主要指网站结构及页面相关性优化，在前两节已经详细讨论过。网站外优化主要就是指外部链接建设，本章将深入讨论。

超文本链接，或者简称为超链接、链接，是互联网的基石。网民很大一部分时间是在浏览网页，而网页与书本报纸上的文章的最大不同之处就在于链接。互联网可以被理解为一个由无数页面所组成的、相互之间交叉连接的巨大网络。以孤岛形式存在的网站和页面是网上的异类，甚至可以说没有什么意义。

早在搜索引擎诞生之前，网民在页面之间穿梭浏览靠的主要就是链接。那时候网站数目很少，站长也都没有商业目的，页面上谈到某个用户可能不太明白的名词时，把这个名词做成链接连向其他有更深入解释的页面，是一个很常态的行为。用户可以通过链接了解新概念，沿着链接无止境地在网上徜徉。站长链接到其他网站的相关资源时没有什么顾虑，不像现在惜链如金。

6.1 外部链接意义

互联网的本质特性之一就是链接。内部链接自己可以控制，在“网站结构”一章已经做了讨论，本章讨论对 SEO 意义更为重大的外部链接。总体上来说外部链接对 SEO 有下面几方面影响。

6.1.1 相关性及锚文字

相关性是搜索结果质量的最重要指标。搜索引擎刚出现时，判断页面与关键词的相关性主要以页面上的元素为基础，也就是页面标题、可见文字、关键词标签、H 标签、黑体等。在某种意义上说，这是根据页面自己说自己与什么相关进行判断的。

按照这种相关性判断方法，只要在页面上多次出现“减肥方法”，甚至堆积关键词，搜索引擎就认为这个页面是与“减肥方法”相关的。

这种相关性和排序算法很快被站长们了解，并且被滥用到令人无法忍受的地步。如果一个页面确实是关于减肥方法的，那么在页面上堆积一点关键词，提高排名倒还有情可原，不过作弊的站长们绝不仅限于此。聪明的作弊者很快想到在页面上堆积搜索次数很高的热门关键词，搜索引擎会误以为这个页面是与热门词相关的，实际上页面很可能

与这些主题毫无关系。站长通过作弊得到搜索流量，用户来到网站后用户体验怎样作弊者就不管了。

看页面自己怎么说的方法很快被证明不可靠，搜索引擎转而看别人怎么说，也就是通过外部链接及锚文字判断相关性。如果很多网站说你的一个页面是在谈减肥方法（以“减肥方法”为锚文字链接至你的网站），尤其如果很多美容健身网站说你的网站是减肥方法领域的专家，那么有很大可能你的网站确实是关于减肥方法的权威。

导入链接内容相关性及锚文字成为判断相关性及排名算法最重要的因素之一，尤其是来自其他网站的导入链接。因为内部链接还是自己说自己，自己吹捧自己不一定可信，别人吹捧就比较可信。如果有很多权威人士吹捧，那么一般就不是吹捧，而是真实的赞扬。

6.1.2 权重及信任度

抛开页面相关性不谈，外部链接使被链接的页面及整个域名权重提高，信任度增加。外部链接越多，发出链接的网站本身权重越高，说明被链接的页面受更多人的信任和尊重。投向一个页面的权重和信任度，也会累计在整个域名上。

权重是现在搜索引擎排名中非常重要的因素之一。除了域名年龄、网站规模、原创性等，形成权重的最重要因素就是外部链接。权重高的域名使网站上所有页面排名能力提高。与被信任网站链接距离近的页面，被当做垃圾内容的可能性大大降低。

权重和信任度与特定关键词或主题没有直接关系。如果你的网站有来自央视、百度、清华大学、华尔街日报这种权重极高网站的链接，你的网站的权重会有质的提升，不管网站目标关键词是什么，排名都会有帮助。

6.1.3 收录

页面收录是排名的基础，不能进入索引库就谈不上排名。

数据观察和经验都表明，外部链接数量及质量对一个域名所能带动收录的总页数有至关重要的影响。没有强有力的外链，仅靠内部结构和原创内容，很难使大型网站收录充分。一个可供参考的经验数字是，PR5 域名大致能带动数百万页面，上千万页面则比较困难。

SEO 们都熟悉的一个链接结构观点是，内页必须距离首页 3～4 次点击之内。原因就在于外部链接在很大程度上决定了搜索引擎爬行深度，一般权重的网站，搜索引擎只会爬行 3～4 层链接。如果达到 PR7 或 PR8，距离首页六七次点击的内页也能收录，从而使整体收录能力提高。

外部链接也是爬行频率的重要决定因素。外部链接越强，搜索引擎蜘蛛重新爬行得越多越频繁，能更快更新内容，发现新页面。权重高的网站，首页几分钟爬行一次都很正常。

由于外部链接对相关性、收录及权重的影响，直接导致关键词排名和搜索流量的不同。现在一般认为，外部链接因素占了 SEO 的 60%～70%以上比重，远超过网站内部优化的重要性。

6.2 Google 炸弹

最能说明外部链接效果的是所谓“Google 炸弹”现象。

1．什么是 Google 炸弹

Google 炸弹（Google Bombing）指的是这样一种情况：

- 数目众多的外部链接指向某一个 URL。
- 这些链接都以特定关键词为链接锚文字。
- 被链接的页面一般并不包含这个关键词，内容和这个关键词基本无关。
- 达到的效果是，这个被链接的 URL 在搜索这个特定关键词时的排名急剧上升，很多时候能排到第一，虽然页面内容与关键词无关。

Google 炸弹大部分是出于恶作剧、政治目的、做实验等目的。Google 炸弹的实现验证了搜索排名算法中的两个事实：

- 外部链接是排名的重要因素之一，一般认为是最重要的因素。
- 锚文字非常重要。

所以当有大量包含特定关键词的外部链接指向某一个网页时，这个网页就算没提到这个关键词，因此没有相关性，排名也会非常好。

2．最有名的 Google 炸弹

美国白宫网站是最富盛名的 Google 炸弹例子。2003 年 10 月，一个叫 George Johnston 的人成功号召一些人用“miserable failure”（惨败）这个关键词为锚文字，链接向美国白宫网站布什总统的个人介绍页（http://www.whitehouse.gov/president/gwbbio.html）。两个月后，在 Google 搜索“miserable failure”，布什的这个个人介绍页升到第一，搜“failure”的时候也是第一。而在这个页面中，不论是 miserable、failure，或是 miserable failure，都没有出现过。

2006 年 9 月，大概白宫有人试图扭转 Google 炸弹所造成的影响，把介绍布什的这个网页做了 JS 转向到介绍所有总统的一个通用页面（http://www.whitehouse.gov/

president/）。结果跳转传输了链接权重和锚文字，这个并不是特别给哪一位总统准备的页面，在搜索“miserable failure”时排到了第一，如图 6-1 所示。

2007 年 1 月，Google 对算法做了修正，防止 Google 炸弹的发生。对具体算法 Google 并没做出解释，也不太可能解释，因为这涉及 Google 算法中非常核心的部分。

众所周知，链接是 Google 排名中最重要的因素之一，哪些链接会被赋予权重，哪些会被降权甚至忽略，Google 当然不会告诉我们。

Google 做了调整后，大部分 Google 炸弹现象消失了。搜索“miserable failure”时，白宫的页面已经找不到了。

图 6-1　2006 年 9 月在 Google 搜索“miserable failure”结果

一般猜想 Google 预防 Google 炸弹的改变有这么几项：

- 短时间内大量外部链接指向某个页面时，有可能是 Google 炸弹。
- 这个页面并没有出现链接中所用的关键词锚文字。
- 关键词锚文字是负面名词。

2007 年 4 月，在 Google 搜索“failure”时，白宫的总统介绍页面再次排到了第一名，不过这次是白宫自找的，因为他们在网页上放上了“failure”这个词：

In a time of war, it’s irresponsible for the… Democratic leadership in Congress to delay for months on end while our troops in combat are waiting for the funds. The bottom line is this: Congress’s **failure** to fund our troops on the front lines will mean that some of our military families could wait longer for their loved ones to return from the front lines. And others could see their loved ones headed back to the war sooner than they need to. That is unacceptable to me, and I believe it is unacceptable to the American people.

这样，本来与 failure 这个词不相关的网页一下子变得相关了，预防 Google 炸弹的算法失去了作用。

2009 年 7 月，还有人继续恶作剧，在 Google 搜索“worst failure ever”（史上最大的失败）时，排在第一的是白宫网站现任总统奥巴马的介绍页面。不过随着网民的讨论，

Google 很快采取了措施。

3．Google 炸弹不限于 Google

2007 年 1 月 Google 算法改变后，负面性质的炸弹消失了。但同样的效应在普通关键词上还是存在的。比如搜索“click here”（点击这里）时，排在第一的是 Adobe Reader 下载页（http://get.adobe.com/reader/）如图 6-2 所示，而这个页面上并不存在“click here”这个词，也没有单独出现“here”，只出现过“clicking”一次。出现这个结果的原因是大量网站用“click here”作链接文字指向 Adobe 页面，推荐人们下载 Adobe Reader。顺便提一下，这个页面是网上罕有的 PR10 页面之一。

同样的道理，排在第三的是 Apple 网站上 QuickTime 下载页（http://www.apple.com/quicktime/download/），页面上既没出现“click”，也没出现“here”。

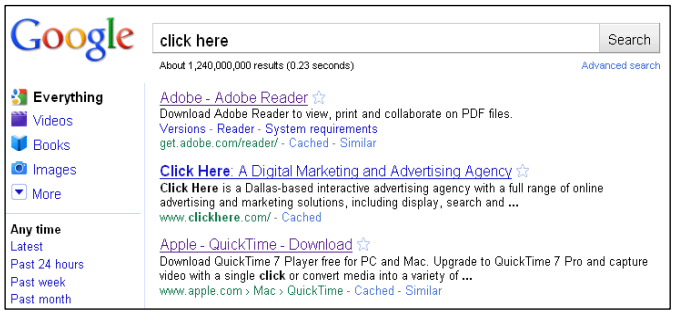


图 6-2　在 Google 搜索“click here”结果

Google 炸弹也不限于 Google。在雅虎搜索“miserable failure”，至少 2007 年年底时排在第一的还是白宫页面。目前（2012 年 2 月）白宫网站小布什总统介绍页面还在必应第七位，百度第一位，如图 6-3 所示。

图 6-3　在百度搜索“miserable failure”结果

6.3　链接分析技术

对 SEO 稍有了解的人都知道链接是网站排名的重要因素，但不一定完整理解链接分析的内容。本节就简单总结链接以哪些方式影响排名。

链接分析技术的含义比 Google PR 要广泛得多。因为 Google 的光环和对 PR 的宣传，很多人把 PR 看得异常重，其实搜索引擎对链接的分析要广泛得多。链接分析包括所有反向链接，不仅限于外部链接。

除了大家耳熟能详的权重传递、锚文字作用，搜索引擎还会分析链接的以下特征。

- 反向链接数目。显然，数目越大，投票越多，对排名越有利。
- 反向链接页面本身的重要性。并不是所有链接都有相同的投票能力，高权重网页的链接对排名影响更大。质量比数量更重要。
- 反向链接增加的速度。增加速度过快，可能引起作弊嫌疑，或进入沙盒。
- 反向链接所在网站的内容主题。来自相关内容网站的链接对排名帮助更大。来自 SEO 博客的链接对美食网站排名不会有什么影响。
- 反向链接所在页的内容是否相关。也属于内容的相关性。SEO 博客里也可能有专门谈美食的一篇帖子，从这篇帖子来的链接对美食网站排名帮助就大一些。
- 反向链接的链接文字，也就是锚文字，是影响网页排名的重要因素之一。
- 反向链接锚文字前后临近的文字。有时候链接文字没有什么意义，比如常见的“点击这里”，链接文字前后的文字可以帮助判断链接目标页的内容。这里所说的“前后”，指的是 HTML 代码中的距离，而不一定是页面显示出来的视觉距离。
- 链接在页面的位置。搜索引擎通过算法可以辨别导航、广告区、页面底部版权声明等区块。链接出现在页面不同位置意味着不同目的。通常出现在正文中的链接才是最有投票意义的链接。
- 外部链接所在域名年龄。历史越长的域名越被信任，来自老域名的链接也更被信任。
- 外部链接所在的域名是否曾经转手。域名所有人一直没有变化，说明网站能持续经营。域名转手后，原来积累的信任度可能会受影响，因为无法保证网站转手后还保持高质量，需要重新考验。
- 反向链接所在页第一次被收录的日期。发出反向链接的页面越老，收录的越早，越被信任。如果这个页面已经存在十几年，比搜索引擎还老，上面的链接显然没有操纵排名的意图，很可能被高度重视。
- 反向链接所在页页面内容是否曾经有变化？有什么样的变化？大部分资料性的网页不会随时间产生明显内容变化，最多是增加更多资料。如果页面内容发生主题方面的重大变化，页面上的链接投票力也会变化，很有可能变得不再内容相关。
- 反向链接第一次出现在页面上是什么时候。一个很老的页面上很早就出现的链接显然有比较高的可信度。最近才出现的链接则需要过一段时间才能走出试用期。链接越老，作用越大，很多做 SEO 的人对此都深有体会。
- 反向链接所在页还链接向哪些其他网站？这些其他网站内容是否相关？质量怎么样？页面上所有链接都指向高质量网站，那么每一个链接的投票力都相应增强，被链接的网站获益也最大。被链接的网站如果内容不相关，整体质量也很低，从这样的页面得到链接，效用不会高到哪里去。

- 外部链接是否有垃圾链接嫌疑？查看一些网站的外部链接，经常能发现绝大部分是来自论坛签名、博客评论，缺少页面正文或博客帖子里的有意义的推荐链接，这绝不是一个健康的链接构成。
- 链接点击率。在搜索引擎能够监测用户行为时，链接的点击率也说明链接的重要性及投票能力。用户的观感更说明问题，用户点击越多，说明对用户帮助越大。
- 用户点击链接后在目标网站停留多长时间。同样也是通过用户行为方式判断网站质量，看到底是否对用户有用。

6.4 什么样的链接是好链接

查看关键词排名时，我们会注意到排名与外部链接绝对数量之间并没有直接对应关系。很多外部链接数目较少的页面会排在有很多外部链接的页面之上。这也是很多 SEO 新手的困惑，为什么自己的页面外部链接很多，排名却不如只有几个链接的竞争对手？这与外部链接的质量有很大关系。

下面从比较理想的情况讨论好的外部链接应该具备哪些条件。

1．点击流量

虽然从 SEO 角度看，外部链接是提高排名的最直接手段，但点击流量才是链接的最初意义。如果能从流量大的网站得到链接，就算链接有 NoFollow 或转向，不能直接提高排名，只要能得到点击流量，就是个好的外部链接，毕竟 SEO 的目的还是得到流量。通过链接直接得到流量，实际上是少绕了一个弯子。

从这个意义上来说，网站 Alexa 排名也可以是快速判断链接质量的标志之一（只要 Alexa 排名不是靠作弊得到的），因为 Alexa 排名能粗略反映流量水平。某些链接带来每天几百几千，甚至上万的点击流量都有可能。

2．单向链接

最好的外部链接是对方站长主动给予的单向链接，不需要链接回去。两个网站互相链接，如友情链接，权重比单向链接要低很多。当然单向链接的获得比交换难得多。正因为难得，价值才更高。

3．自发及编辑

好的链接是对方站长自愿自发提供的，而且通常是站长的编辑行为，也就是说在文章中提到某个概念时，认为你的页面是最好、最权威的，所以链接到你的页面。这种有编辑意义的链接，才是真正意义上的投票。

4．内容相关性

寻找外部链接来源时，内容相关性非常重要。一群 SEO 人员说某个页面是关于减肥方法的，不一定可信；一群健身教练说某个页面是关于减肥方法的就更可信，投票力也越强。

内容相关度既适用于整个网站级别，也适用于页面级别。网站主题相关当然最好，有的时候某个页面不一定与整站主题完全吻合，页面本身的主题与被链接页面相关，也比无关页面的链接权重高。

内容相关度的判断并没有明确界限，只能靠外链建设人员的常识和直觉。我们无法明确知道美容、健身、医药主题到底哪个与减肥最相关，但是至少可以知道这些主题比金融股票与减肥的相关度要大。

5．锚文字

锚文字中出现目标关键词是最好的外部链接，在搜索引擎排名算法中有很大比重。但锚文字也不能过度集中。一个网站首页获得的外部链接全都使用一个锚文字，通常就是首页的最重要目标关键词，往往是导致惩罚的原因之一，因为太不自然，刻意优化的痕迹太明显。

所以在可能的情况下，来自重要页面的链接尽量使用目标关键词做锚文字。权重不太高的页面，适当混合各种各样的锚文字。自发获得的链接通常锚文字五花八门，你无法控制别的站长用什么链接文字，从根本上避免了锚文字过度集中的问题。

6．链接位置

页脚、左侧或右侧导航中的广告部分，专门设置的友情链接页面都是最常见的买卖链接和交换链接位置。搜索引擎通过对页面分块，可以鉴别出这些位置的链接，降低其投票权重。最好的外部链接出现在正文中，因为只有在正文中才是最有可能带有编辑意义的自发链接。

7．域名权重及排名

发出链接的域名注册时间、PR 值及网站首页目标关键词排名如何，都直接影响链接的效果。寻找链接来源时，最好搜索一下对方网站首页目标关键词（通常在首页 Title 中有体现），也要搜索一下对方网站或公司名称，看看排名在什么地方。排名越好，说明对方网站权重越高。搜索对方目标关键词并不一定要求排名在前十或前二十。就外部链接获得来说，排名在前十几页都说明有一定的权重和排名能力，是不错的链接来源。

8．页面权重及排名

除了网站整体权重和排名，发出链接的页面权重及排名能力也是需要注意的地方。

除了使用首页交换链接外，从其他网站首页获得链接是很困难的，从内页获得链接就容易得多。

有的时候，从内页发出的链接不一定就比首页链接效果差多少。很多网站内页本身就有大量外部链接，内页的权重也很高。可以使用 SEOMoz 的网站最强页面工具（见 SEO 工具一章）检查某个域名上有哪些内页权重也不错，获得首页链接难度大时，可以主动请求从权重高的内页获得链接。

9．导出链接数目

页面上导出链接越多，每个链接所能分得的权重就越少。很多专门用于友情链接交换的页面没有其他实质内容，全部是导出链接，可能多达几十上百个。从这样的页面获得链接效果就很差了。有实质内容的博客帖子、新闻页面等，通常导出链接要少得多。

10．页面更新及快照

如果页面在搜索结果中的快照很新，说明这个页面经常被搜索引擎蜘蛛重新抓取。这样的页面不仅权重、投票力比较高，也意味着加上去的链接能很快被检测到，计入排名算法。如果页面快照是几个月前的，获得链接后可能还要过几个月才能被收录。

11．网站整体健康情况

网站首页 PR 值？主要分类页面是否获得相应的比首页低一级的 PR 值？是否被开放目录、雅虎等收录？网站规模？主要关键词及长尾关键词排名情况？是否有被惩罚的迹象？

12．来自好邻居

寻找外部链接时，只关注正规网站，不必考虑垃圾网站、色情赌博等违法内容网站。网站被搜索引擎删除或严重惩罚也不能考虑。尤其在交换链接时，这种网站都要排除在外，因为链接到这种网站，对自己的网站一定有负面影响。

就算不存在交换，对方网站单向链接过来，也不会有好的效果。如果这样的链接过多，还可能被搜索引擎认为你的网站与这些有问题的网站为邻，导致惩罚。

13．来自 edu，gov 等域名

大部分 SEO 认为，来自 edu，gov 这种不能随便注册的域名的链接效果最好。Google 工程师多次否认这一点，他们的官方说法是，这些域名并不天生比其他域名权重高。

实际情况是，通常 edu，gov 域名上垃圾内容比较少。这些网站由于与其他政府、教学、科研机构的关系，本身获得高质量外部链接的机会比商业网站高得多。所以就算 edu，gov 域名并没有天生的优势，却往往有累计下来的质量优势，他们给出的链接效果一般也更好。

当然，上面讨论的是最佳状况。如果这些条件都能拿到高分，这样的外部链接可以说是极品，可遇而不可求。做外部链接建设时也不可能拘泥于这么高的要求，只能尽量接近而已。

6.5 外部链接查询

做 SEO 就肯定要经常查询某个网页或某个网站的外部链接，不光是查自己的网站，也会查竞争对手的网站。

6.5.1 链接查询指令

查反向链接和外部链接最简单直接的方法是在搜索引擎中使用高级指令搜索。也有一些线上工具帮助自动查询，但也是向搜索引擎发同样的查询指令，然后整理显示在页面上，与自己人工查询效果是一样的。下面对几大搜索引擎的反向链接查询指令做个比较总结。

1．Google

Google 反向链接查询指令是：

link:www.chinamyhosting.com

查询页面外部链接则是：

link:www.chinamyhosting.com -site:chinamyhosting.com

不过，Google 给出的反向链接数是最没用的，基本上可以忽略。Google 显示的反向链接数只是 Google 知道的一部分，更新的时间表只有他们自己知道，列出的反向链接顺序也没有什么规律。所以任何基于反向链接的工作，都不应该以 Google 反向链接查询为准。

2004 年前 Google link:指令返回的反向链接是 PR4 以上的网页，但是现在 Google 返回的反向链接基本上没有任何规律，近乎是随机的一个子集。

Google 不列出所有反向链接，也不按重要性排列，恰恰是为了防止 SEO 们研究出反向链接与排名的对应关系，也防止黑帽们给重要页面制造垃圾。

2．雅虎 Yahoo!

Yahoo!的反向链接查询指令是：

link:www.chinamyhosting.com

或

link:http://www.chinamyhosting.com

显示的是指定 URL 的反向链接。

另一个有用的指令是：

linkdomain:chinamyhosting.com

返回的是整个域名的反向链接。

linkdomain:chinamyhosting.com -site:chinamyhosting.com

这个指令给出的就是整个域名的外部链接。

雅虎给出的反向链接数据是最准确的。虽然雅虎给的数据也不一定十全十美，但对大部分与反向链接有关的工作是最有指导意义的。雅虎列出的反向链接大致是按反向链接页面的重要性排列的。

3．微软必应 Bing

早在微软的搜索引擎还称为 MSN Search 时也曾经支持 link：指令，但目前 Bing 不支持 link：或 linkdomain：指令，所以无法查询 Bing 收录的网站反向链接。

不过 Bing 有一个其他搜索引擎都不提供的指令：

linkfromdomain:chinamyhosting.com

显示的是 chinamyhosting.com 连向其他域名的所有链接。

4．百度

就我所知，百度也没有指令用于查询反向链接。

6.5.2 工具查询外链

使用工具查询外部链接在本书其他章节中有更详细的介绍，这里只简单总结。

Yahoo! Site Explorer：最常用的外链查询工具，详见 SEO 工具部分第 11～12 节。Yahoo! Site Explorer 与 linkdomain：指令显示的数据相近，但不完全相同。

Google 网管工具：详见竞争研究部分 3.12.4 节。数据比较完整准确，可惜只能看到自己网站的链接数据。

OpenSiteExplorer：详见 SEO 工具部分第 11～27 节。SEO Moz 开发的工具，基于自己爬行建立的链接数据库。

Majestic SEO：详见 SEO 工具部分第 11～32 节。不是显示搜索引擎收录的外链，而是 Majestic SEO 自己抓取、建立的数据库。非常详尽的外链数据。

6.5.3 影响排名的链接

使用指令或工具查询得到的链接数据中有几个数字值得特殊关注：总链接数、总域名数和高质量链接数。

总链接数顾名思义就是外部链接总数量。

总域名数指的是外部链接来自于多少个独立域名。通常总域名数远远小于总链接数，因为有的外部链接是全站链接，大的网站在页脚上全站链接到你的网站，可以带来成千上万个外部链接，但只是来自于一个域名。

根据 SEOMoz 的数据研究，总域名数与排名有更强的关系，其作用比总链接数更大。因此，来自 100 个普通权重域名的链接，比来自 1 个权重较高的域名上的 100 个链接作用更大。现在有不少网站出售链接，PR 值高的链接价格相当高。与其每个月花几千块买一个 PR7、PR8 的链接，还不如把同样的预算花在来自不同域名的几百个 PR4、PR5 链接上。

SEO Moz 研究针对的是 Google 排名，不一定适用于百度和其他搜索引擎。不过在百度做一下简单的调查，也可以发现同样的规律。表 6-1 是 2010 年 4 月在百度搜索“减肥方法”，排在前十名网站的链接数据。已经排除了百度百科、百度知道，因为百度自己的内容有特殊权重，不足以说明问题。

表 6-1 中列出了根据 Majestic SEO、OpenSiteExplorer 两个工具查询的总域名数和总链接数，以及根据雅虎 Site Explorer 得到的总链接数。可以看到，除了新浪等权重极高的网站总域名数不高排名却很高，说明域名权重是排名算法中的重要因素之一，其他页面排名大致与总域名数有对应关系，总域名数越大，排名越好。感兴趣的读者可以做更大规模的统计。

表 6-1 百度搜索“减肥方法”排名页面链接统计

页　面	Majestic SEO 总域名数	Majestic SEO 总链接数	OpenSiteExplorer 总域名数	OpenSiteExplorer 总链接数	Yahoo SiteExplorer 总链接数	Google PR 值
www.ssnn.net/	16 000	157 106	4305	23 589	56 600	6
eladies.sina.com.cn/fat/	290	66 574	56	3419	31 822	7
www.top002.com	3550	31 940	495	5145	6418	4

续表

页　　面	Majestic SEO 总域名数	Majestic SEO 总链接数	OpenSiteExplorer 总域名数	OpenSiteExplorer 总链接数	Yahoo SiteExplorer 总链接数	Google PR 值
www.di5j.cn	6703	144 558	1298	21 429	32 820	6
lady.163.com/fit/	1018	5474	150	2915	2686	6
www.lady8844.com/shoushen/	1727	20 034	319	28 924	9894	5
fitness.pclady.com.cn/jf/	1610	8712	294	3522	2935	5
fitness.39.net/jfff/	150	346	20	209	93	4
lady.163.com/special/002635SE/soprtsfitting.html	230	814	35	197	231	4
www.fx120.net/fitness/	226	998	73	24 169	762	5

在做外部链接时，如果看到竞争对手有几万几十万总链接数，先不要灰心。查看一下对方的链接总域名数，很可能不超过几百或几千个，要追上竞争对手还是有希望的。链接总域名数达到几千个就已经很不错了。

另外一个影响排名的链接数字是高质量链接数。无论用哪种工具查询，如果排在前面的链接都是论坛博客留言，说明这个竞争对手并不可怕，哪怕总链接数很高。

6.6　外部链接原则

在讨论具体外部链接方法之前，我们先来探讨建设外部链接时需要遵循的几个原则。

6.6.1　难度越大，价值越高

实际操作过 SEO 的人都知道，原创内容和外部链接是两个难点，都是费时费力的工作。外部链接甚至无法保证有投入就能有产出，获得好的外部链接就更难了。

不过一般来说，越是难度大的链接，效果越好。SEO 人员千万不能因为第一次联系时被拒绝就灰心丧气而放弃。很多时候从权重高的博客、新闻网站、论坛获得链接，需要与对方站长联系很多次。有时不能上来就要求链接，还要先与对方交朋友、互相帮助，有了一定交情后要求链接，才能变得自然而然。

一些权重高的网站并不接受友情链接，只有对方了解你、相信你之后，才可能给你一个单向链接。这个过程也许要花上几个月时间。越是这种难以获得的链接，才越有效果。

6.6.2 内容是根本

“内容为王，链接为后”是 SEO 行业的老生常谈。有一些页面就内容质量来说并不是最好的，但因为链接的关系排名最好，所以有的 SEO 人员认为链接才是最重要的。不过我依然认为内容为王，链接为后，原因在于高质量的内容可以带来高质量的链接，而链接不能使你的网站产生内容，所以最后其实无法获得高质量的外部链接。

就像搜索引擎不在乎你的网站一样，其他站长也不在乎你的网站。但站长都在乎自己网站的用户。要想让对方链接到你的网站，你必须为对方网站用户提供价值，最重要的价值就是高质量内容。站长在自己网站上不能提供足够内容时，才会以外部链接的形式导向其他提供相关内容的网站。天下没有白吃的午餐，没有高质量的内容，获得的链接就只能是交换、购买或垃圾链接。

6.6.3 内容相关性

这一点前面已经讨论过，寻找外部链接时，内容相关性是最重要的考量标准之一。正因为如此，参与相关论坛的讨论，访问同行业内其他人的博客并留言，相互沟通和支持，成为行业内的积极参与者，对一个 SEO 人员来说是非常重要的。

6.6.4 链接来源广泛

上一节讨论了什么样的外部链接是好链接，但一个正常的网站不可能全都是好的链接，而没有一般的、甚至质量比较差的链接。进行外部链接建设时，也应该大致上使外部链接构成自然、随机，来源广泛，呈现出健康正常的特征。

这里所说的来源包括：

（1）网站种类。既有博客链接，也有论坛签名，又有新闻网站、社会化媒体、商业网站等。

（2）链接位置。既有出现在页脚、导航条的链接，也有正文中的链接。前面说过页面正文中的链接效果最好，但一个网站的外部链接全都处在正文中也不正常，操作痕迹太明显。全是页脚的链接就更不像是对用户有帮助的链接。

（3）各种 PR 值的链接都有。正常的网站一定有来自高 PR、高权重的链接，也有来自低 PR，甚至没有 PR 值页面的链接。如果一个网站的外部链接全都来自高 PR 值的页面，就显得很可疑，按常理判断，恐怕以购买链接居多。

（4）不同域名。尽可能从各种不同域名获得链接。.com，.net，.org，也有政府、学校网站，还可以有不同国家域名的链接。总域名数越多越好。

（5）IP 地址。只要不是链接工厂，正常外部链接建设得到的结果，一定是链接来自大量分散的 IP 地址，之间没有什么关系。如果大部分链接来自几个特定 IP 地址，被怀疑为链接工厂的可能性将大大提高。

（6）新旧网站。和 PR 值一样，外部链接应该来自各种历史的网站，既有新的，也有旧的。

6.6.5 深度链接

购买链接、交换链接一般仅限于首页及几个重要分类页面，很难顾及到网站上的大量内页。一个靠高质量内容吸引外部链接的网站，则不仅首页有链接，还能自然吸引到连至内页的深度链接。

深度链接不仅使外部链接构成趋向自然，对内页权重也有很大影响。我注意到一个现象，很多网站首页 PR 值不错，按正常情况，一级分类页面 PR 值和权重应该只比首页差一个等级。实际上却不一定这样，很多一级分类页面 PR 值为零，从排名能力看，这些分类页面似乎并没有全部接收到内部链接传递过来的权重。只有当这些分类页面本身有外部链接时，PR 值、排名能力和权重才提升到应有的位置。也就是说没有外部链接的内页，似乎连内部链接的权重也不能完全体现出来，有了外部链接，内页的内部链接权重及排名能力才完全释放出来。

6.6.6 锚文字分散自然

前面说过锚文字对页面相关性影响很大，以目标关键词作为锚文字效果最好。但锚文字高度集中又常常是被惩罚的原因，所以锚文字的构成也必须自然而然。就像我的博客外部链接锚文字各种各样五花八门，既有使用“SEO 每天一贴”的，也有使用“ZAC 博客”的，链接到内页（具体帖子）的链接锚文字则更分散。

6.6.7 平稳持续增加

外部链接建设最忌讳突击，花一个月时间增加大量外部链接，看到效果后却不再持续增加。真正对用户有用的网站，外部链接都是随时间平稳增加的，很少大起大落。没有新链接也经常是排名稳步下降的原因。与其一个月增加 100 个链接，然后几个月没有动静，不如把时间精力分配到几个月时间里，每个月增加二三十个，并坚持下去。

6.6.8 质量高于数量

外部链接毕竟只是排名的因素之一，虽然可能是最重要的因素。外部链接绝对数量与排名并不呈线性关系。一个高质量外部链接常常比几十几百个低质量链接有效得多。这在 Google 排名中更明显。

统计一下 Google 排名前两页的页面外部链接数量就可以看到，排名位置与链接绝对数量几乎呈现随机关系，看不出明显关联。但是两三个高质量链接常常使排名有质的飞跃。这也是很多 SEO 查看排名时疑惑的地方，有的页面从链接数量看没什么值得称道的，排名却稳稳处于前列——我们不一定能发现质量最高、真正起作用的那几个链接。

这里所说的外部链接原则实际上针对的是刻意的人工外部链接建设，对新站、小站来说也是不得已而为之。如果是真正以高质量原创内容取胜的网站，人工外链建设甚至可能是不必要的，其他站长自愿自发给予的链接数量就已经足够了，而且完全符合上面讨论的原则。

6.7 网站目录提交

网站目录编辑审核站长提交的网站，按一定的分类方法把收录的网站放在适当的目录分类下。网站目录并不抓取网站上的页面，只记录下网站的网址、标题、说明等。网站目录也常称为网址站、导航站等。

提交和登录网站目录是早期常用的网站推广手法。在搜索引擎出现之前，网上用户大多是通过网站目录寻找网站的，如雅虎、开放目录。随着搜索引擎的发展和被普遍接受，现在网站目录对用户的重要性越来越低了。在英文网站领域，已经很少有人通过网站目录来寻找要访问的网站。只有几个最重要的网站目录，还能带来一点点直接点击流量，比如雅虎、开放目录、Business.com。

被高质量的目录收录对 SEO 依然有重要意义，因为带来不错的外部链接。所以寻找、提交网站到网站目录，现在是 SEO 人员必做的功课。

中文网站目录比英文的更有价值，有不少历史悠久、口碑好的网址站本身流量十分巨大，很多刚刚接触互联网的新手把网址站作为自己访问其他网站的出发点。很多网吧还把网站目录设置为浏览器首页。最著名的中文网站目录 hao123，是流量最大的中文网站之一，无数用户先到 hao123，再点击去其他自己感兴趣的网站。

各个网站目录收录标准不同，有的付费就行，有的给交换链接就行，也有的对提交网站要求比较高。下面以最难进入、要求最高的标准（以开放目录为典型）讨论目录提交，要求较低或付费能解决的目录更迎刃而解了。

6.7.1 提交前的准备

在提交之前，首先确保自己的网站有可能被网站目录收录，包括：

（1）内容的原创为主。高质量网站目录不会收录那些粗制滥造，完全以采集、抄袭拼凑而成的网站。只有原创内容丰富的网站，才能给网站目录本身带来价值。

（2）网站已经全部完成。不能出现 404 错误、打不开的链接、显示不出来的图片、“网站正在建设中”之类的文字。确保整个网站已经完成，所有功能正常运行。

（3）页面设计达到专业水准。与搜索引擎抓取页面不同，目录是由编辑审查提交的网站。网站设计给编辑的第一印象十分重要，如果设计过于简陋或业余，内容再好，编辑也很可能没有心情仔细审查。

（4）联系方式齐备。网站上应该清楚地标明公司或站长的联系方式，包括电子邮件地址、通信地址、电话。这显示着网站的正规和专业性。一些高质量的网站目录，如开放目录，对此有硬性的规定，凡是没有通信地址和电话的，一般情况下不予收录。

事先撰写好提交过程中可能需要用到的三个信息：

（1）网站标题。通常标题就应该是网站的官方名称。在可能的情况下，可以适当加进一些关键词，但是绝不要因为要加关键词把标题写得广告性太强。越是正规、质量高的目录，越是应该使用网站官方名称，哪怕名称中完全没有关键词。

切记不要在标题中加入口号式、宣传式、广告式的语言。如公司网站是房地产行业，就把标题写为“鲲鹏房地产公司”，而不要写成“最专业的房地产公司”。

（2）网站说明。用一到两句话简要说明网站的内容和功能。同样，切忌在说明中使用广告性语言，诸如“最好”、“最便宜”之类的文字，目录编辑对自吹自擂的语言有天生的反感。只要用平实的第三方角度简要叙述出网站的内容就可以了。

（3）关键词。有的网站目录还允许提交关键词，方便目录站内搜索使用。选出与网站最相关的，可能被搜索次数最多的 3～5 个关键词。

标题、说明及关键词事先都要准备好。虽然有的网站目录不一定需要所有这三部分内容，但是事先花点时间撰写好备用，提交时会节省很多时间。

6.7.2 寻找网站目录

怎样找到能提交的网站目录呢？

首先最简单也最有效的方法，就是在百度或 Google 搜索与网站目录相关的关键词。虽然本书中都是使用“网站目录”这个词，但可以搜索的关键词还有很多，包括网站目录、目录提交、目录登录、分类目录、网址提交、网站登录、网址站、网站大全、导航站、网站导航等。

搜索这类词返回的网站数目巨大，很可能在几百万甚至上千万个结果以上。为缩小目标，还可以在搜索词中加上自己网站的关键词或地理位置，如房地产目录，儿童服装

网址站，北京网址大全等，这样得到的搜索结果更有针对性，可以首先着手提交。利基市场或地域性目录都提交完了，再去提交比较宽泛的网站目录。

第二个方法是看竞争对手都在哪些目录中被收录。这可以通过查询竞争对手外部链接找到。

第三个简单方式是，很多网站目录就收录有其他网站目录和网址站，尤其是与站长或网站建设、网络营销相关的网站目录。所以找到一个网站目录，就可以顺藤摸瓜，找出一串可以提交的网站目录。如开放目录（dmoz.org）收录的分类目录：

Top：World：Chinese Simplified：计算机：互联网络：搜寻：分类目录

寻找网站目录时需要大致判断目录质量，尤其是对方要求友情链接或付费时。几个简易又重要的指标是：

- 首页 PR 值。
- 内部分类页面是否有 PR 值。
- 已收录的是些什么网站？如果目录充斥垃圾网站，就不要在这里浪费时间。

6.7.3 网站提交

找到要提交的网站目录后，还要正确选择向哪一个分类提交网站。

网站目录都是按特定的方式进行分类的，提交时一定要在与自己网站最相关的那个分类中提交。有的站长喜欢把网站提交到比较大、层次比较高的分类中，实际上被收录的机会反而更小，应该一直找到最适合的小类里提交。

如果不很确定应该在哪个分类提交，可以搜索一下主要竞争对手是在哪个分类中收录的，就到哪个分类提交。

找到最适合提交的类别后，通常页面上有一个提交网址的链接，在提交页面上填写事先准备好的标题、说明、关键词，当然还有最重要的网站 URL。提交表格后，就只能耐心等待了。

现在很多中文网站目录要求提交网站做一个友情链接，才会批准收录这个网站。这时候站长就应该自己做一个决定，是花更多时间寻找到那些不需要做友情链接的网站目录，还是做友情链接？要求友情链接的网站目录实际上和交换链接没有太大区别。

建议站长在提交要求友情链接的目录时，只选择那些相关性高或质量比较高的目录。如果碰到一个网站目录不管相关与否都提交，都做友情链接，那么要做的友情链接就太多了。因为不相关的友情链接而放弃网站目录一点儿都不可惜。到网上逛一逛，能找到太多的网站目录。只要花时间，其实还是可以找到不需要友情链接的目录，比如一些行

业网站、地方性网站等。

提交过的所有目录都要做个记录，包括目录地址、提交时间、被收录时间、被收录的具体类别，自己网站上的友情链接页面等。目录提交是一个长期的、烦琐的过程，如果没有记录，时间久了就很难记得自己的网站已经提交过哪些目录，在哪里被收录，哪里一直没有回音。

提交网站后一两个月如果没有收到对方回信，网站也没有在相应类别中出现，可以再提交一次。如果还是没有消息，也不必太过执着，放弃这个目录，去寻找其他目录就可以了。

这里所说的目录选择和提交过程实际上是以质量比较高的网站目录为目标的，很多小型目录并没有这么高的要求。比如说可能允许在网站标题中堆积一些关键词，只要你和他交换友情链接，对方就收录。高质量的、要求比较严格的网站目录如果都能正确提交和收录，其他那些要求不高的目录就更容易处理了。

6.8 友情链接

友情链接，或叫交换链接，是外链建设最简单也最常见的形式。我链接向你，你链接向我，互相给对方带来一定的点击流量，也有助于搜索排名。

中文网站交换友情链接比英文网站要普遍得多。正规的英文网站就算交换也很少会在首页交换友情链接，通常是开设一个交换链接部分，把友情链接都放在专用的友情链接页面上。中文网站不仅有友情链接页，大部分网站还接受首页友情链接，连很多门户和大公司网站也是如此。友情链接在中文网站推广中是个常态，站长们十分熟悉，搜索引擎也不得不考虑到这一点。

有的 SEO 人员认为交换链接没什么用，已经被搜索引擎大幅度降权。我的观察和看法是，来自主题相关的、正规网站的链接还有不错的效果，尤其是中文网站。实际上，友情链接往往是无法避免的，没有刻意交换的网站也如此。我的博客没有交换友情链接，但 blogroll 中列出我确实在读的几个中文 SEO 有关博客，这些博客有一部分也在读我的博客，所以在他们的网站上也有链接指向我的博客。虽然我们没有交换，但形成了事实上的交换链接。

6.8.1 友情链接页面

除了首页外，友情链接也可以放在内页，常见的是两种形式。

（1）在网站上开设专门用于交换友情链接的部分。如果只计划小规模交换友情链接，这部分可以只是一个页面。计划大规模交换时，可以按主题进行分类，把友情链接放在不同主题页面上。在策划网站框架时就应该根据网站自身内容，按相关主题把友情链接

分成 10～20 个页面。如果站长野心更大，网站规模也更大，可以分成更多类。整个友情链接部分类似于一个小型网站目录。

这样做也有它的缺点。很多站长会觉得友情链接放在首页上最好，做成小型目录形式，友情链接大部分需要放在距离首页一两次点击的主题页面上。不过如果网站结构合理，这些友情链接主题页面同样可以得到很好的收录及不错的 PR 值。

（2）在正常分类页面上留出友情链接位置，把友情链接直接加在分类页面上。这样友情链接成为网站自然的一部分。

将友情链接放在多个内页，长远来看更有扩展性。首页的位置终归是有限的，不可能放上几百个友情链接。一个友情链接页面能放的也有限，如果真的把一两百个友情链接放在一个页面上，给予每个链接的权重将大大降低，站长们很可能不愿意与这样的网站交换。

6.8.2 软件使用

有一些现成的软件可以帮助管理专门设立的友情链接页面。站长在软件后台创设新分类，也可以人工添加友情链接，其他站长可以在友情链接页面上自行提交友情链接申请，站长在后台检查对方是否已做好连过来的友情链接，并进行网站标题和描述的审核和批准。

软件也会定期自动检查已经批准的友情链接，看对方网站是否还保留着连过来的友情链接。如果对方因为某种原因已经拿掉连回来的链接，站长在后台会看到提示，可以进行人工审查及进一步处理。

虽然有软件辅助省时省力，不过建议最好不要使用任何软件，尤其是不要使用网上大家常见、常用的现成软件。友情链接管理软件生成的页面往往相似度太高，在代码、页面排版、文字措词、分类等方面有明显的痕迹。如果你的网站和其他成千上万网站有特别高的相似度，用户不喜欢，搜索引擎同样也不喜欢。有证据显示，几个常见的英文友情链接管理软件已经被搜索引擎检测和惩罚。所以还是建议，要么完全人工管理友情链接页面，要么使用自己写的管理软件。

6.8.3 寻找交换链接目标

寻找友情链接伙伴相对简单，网上有很多渠道。

- 在搜索引擎搜索目标关键词加“友情链接”、“交换链接”等词，会看到很多接受友情链接交换的相关网站。
- QQ 有不少站长、交换链接群。

- 站长聚集的论坛如 admin5.com，im286.com，chinaz.com，都有很活跃的链接交换版块。
- 还有一些链接交换专用平台，如 go9go.cn。

判断是否合适交换链接先看对方网站年龄。比较老的网站可信度更高，今后还将继续存在的可能性也更大。

对新网站则需要注意看发展潜力如何？查看一下 Alexa 排名是否稳步提升中？Google 趋势显示的流量是否增长中？网站是否持续更新？站长是否在用心做站？很多新网站有较大的发展潜力，这时候你找他交换链接成功的机会更大。一旦新网站过几年变成一个成功的大网站，想成为友情链接合作伙伴，人家该挑挑拣拣，不一定看得上你的网站了。

很多站长在寻找友情链接伙伴时会特别关注 PR 值。当然这是一个可以参考的指标。如果一个很老的域名首页 PR 值却一直是零，这多少有些可疑，有可能是被搜索引擎惩罚，有可能是对方站长从没有认真推广过网站，以后变得认真起来的概率也不大。但 PR 值不是唯一决定性因素。新网站 PR 值为零很正常，只要有发展潜力，不妨考虑。

寻找友情链接时还要注意网站内容的相关性。虽然友情链接按说应该是从友情出发的，但站在用户角度考虑，就算两个站长真的有友情，如果一个是 IT 资讯网站，一个是育婴网站，用户点击友情链接的可能性将大大降低。用户觉得没有用的东西，就是对网站没帮助的东西。交换友情链接最重要的考量之一是看能否带来有效流量，对方网站流量大，用户活跃，内容相关性又高，才是最佳选择。

6.8.4 交换链接步骤

如果有对方站长 QQ 或 MSN，当然沟通起来很方便，下面讨论的要点都能实时反馈，迅速解决，大大简化。主动联系时注意基本的礼貌，被拒绝也不要计较，把时间花在更有效率的事情——寻找下一个目标上。

下面就以最麻烦的邮件沟通为例讨论一下需要注意的地方。做英文网站通常只能邮件联系或填写在线表格，欧美站长很少有上 IM 的习惯。

在发邮件与对方联系之前，应该先把对方链接放在自己的网站上，这是个基本的礼仪。我相信所有站长都收到过交换链接请求，邮件里说希望和你交换链接，只要你把他的链接放上，他就链接回来。我想大部分站长看到这样的邮件直接就删除了。当你首先联系其他人时，先把自己该做的做到，放上人家的链接，不要奢望你找人家办事，却希望人家先链接向你。

发链接请求邮件前最好看看对方是否接受友情链接。如果对方网站上既没有首页友情链接，也没有专用的交换链接页面，网站上也没有可能是交换的“合作伙伴”之类链

接，就不要以友情链接的名义联系。对方很明显接受和欢迎友情链接时，通常会列出交换链接步骤，比如填写在线表格或发邮件时应该包含哪些内容，尽量按对方列出的步骤和要求去做。

邮件中用一句话写清自己和对方网站的基本情况，最起码得说清是哪个网站。我经常收到交换链接邮件，里面根本没提是想和我的哪个网站交换链接。大部分站长运营不止一个网站，接到这样的邮件就明白对方根本不是诚心诚意，大概只是拿一个邮件模板在群发而已。

邮件中最好表明你真的浏览了对方网站，并且觉得双方网站能够良好互补，所以希望交换友情链接。还要告诉对方，已经把对方链接放在哪一个页面上，欢迎对方来检查。

提供你希望对方使用的链接文字及简短说明，对方站长同意交换链接时可以参考使用。甚至可以写好链接的 HTML 代码，既让对方省事，也对链接文字多些控制。

发出邮件后两到四个星期如果没有收到对方的回复，可以再发一封邮件提醒一下，但是绝不要有任何威胁或不满的口气。第二封邮件只要写个友好的提醒，并且向对方表示，不接受这个友情链接也可以理解，以后有机会再合作。

发出提醒邮件后如果还没有消息，就不要再提醒或催促对方了。如果人家愿意交换，已经这么做了。人家不愿意，千万不要强求，那是浪费自己的时间。继续下一个目标就可以了。

交换链接是一个长期又烦琐的工作，但效果也是明显的。持之以恒，必有收效。

6.8.5 内页正文链接交换

要达到最好的友情链接效果，可以考虑内页正文中的链接交换，也就是不要把链接放在通常放友情链接的地方（专设的友情链接页面，页脚或侧栏中的友情链接位置），而是放在普通信息页面（帖子、文章、新闻等）的正文中。

前面讨论过，正文中的链接价值比较大，被判断为友情链接的机会也比较小。置于正文中的友情链接与自然链接很难区分。

当然，要在内页正文中交换链接，要做的准备工作更多：

- 找出对方权重较高、比较可能给链接、内容又相关的内页。
- 建议对方在原有内容的哪处文字加上链接。如果能以目标关键词为锚文字，效果更好。
- 自己网站上哪个内页的什么地方适合给对方做链接。
- 可能的话，双方链接不要同时加上。

过程要烦琐得多，但如果能从一些权重高的相关文章页获得正文中的链接，效果比普通友情链接好得多。

6.8.6 交换链接中的小花招

交换友情链接时，大部分站长还是能认认真真、老老实实，双方诚信地合作的。也有的站长喜欢耍点小花招，试图欺骗性地得到友情链接却不给予相应的回报。下面举一些例子，读者以后也能有所提防。

1．交换完链接后再删除链接

最简单的花招就是交换完链接后，过一段时间悄悄把链接拿下来，这样你连过去的链接就成了单向链接。

这种还算容易发现。有的站长用程序自动检查对方链接，有的不使用程序，但也应该经常看看友情链接伙伴网站上你的链接还在不在。如果对方悄悄拿下了，我个人认为也没必要问对方是怎么回事，你也别客气，把他的链接也拿下来就行了，以后再也不必相信这个站长了。对一些确实很有价值并且是大公司网站的友情链接，可以考虑询问一下原因，有可能是因为管理混乱、人员交接等出现误删。小公司和个人站长不小心删除友情链接的可能性就很小了。

2．刻意把友情链接页的权重降低

设有专门的友情链接页时，有的站长通过控制站内链接结构使友情链接页面得到的权重很低。

最典型的方法是只在首页或网站地图放上友情链接页的链接，其他页面都没有连向友情链接页的链接，或者链接放上 nofollow 属性。这样，友情链接页可以被搜索引擎收录，但整个网站只有一两个页面链接向友情链接页。这样的链接结构使得友情链接页的权重非常低。

正常的网站结构应该使友情链接页成为整个网站的有机组成部分之一，所有处理方法和其他页面相同。比如说友情链接部分就当做一个栏目处理，里面又分很多内页。这个栏目和其他内容栏目一样，在所有页面上都有一个导航链接（出于用户体验，可以放在页脚），使友情链接页面得到应有的权重。

3．使友情链接页根本不能收录

有的站长使友情链接页看似普通网页，链接结构也正常，但其实使用 robots.txt 文件或 meta noindex 标签使友情链接页根本不能被搜索引擎收录。这是我建议交换链接时还得考虑链接页 Google PR 值的原因之一。页面有 PR，至少说明可以被搜索引擎收录。

4．友情链接本身不传递权重

有的站长给友情链接加上 nofollow 属性，有的在页面 HTML 代码头部加上 meta nofollow 标签，使页面上所有链接不能传递权重，不看源代码不容易发现。有的使用脚本转向，这样的链接实际上已经不是正常链接了，一般也不能传递权重。

有的站长做得更隐蔽一些，“链接”是经过转向的，但通过使用脚本，鼠标放在链接上时浏览器状态栏却显示正常的链接 URL。这样，除非你去检查页面的源代码，否则很难发现对方给的链接其实是通过脚本转向的。

比如，下面的链接代码：

```
<a href="http://www.domain.com/redirect.php?partner"
onMouseOver="window. status='http://www.partner.com';
return true" onMouseOut="window.status=''">Partner</a>
```

链接其实是通过 redirect.php 转向的，但鼠标放在链接上时，浏览器状态栏显示的目标 URL 是 http://www.partner.com。

5．链接页可能根本就是只给你准备的

有的站长心机更重，在发友情链接交换邮件时告诉你，你的网站链接已经放在比如 http://www.domain.com/index.php 页上。点过去一看，果然有你的链接，而且还是首页，你就链接回去了。

但如果你再仔细检查一下对方网站，会发现网站首页根本不是你看到的这个页面。去掉 index.php 文件名，访问网站 http://www.domain.com 时，真正的首页是另外一个页面，有可能是 index.html 文件。

通常服务器配置使 html 文件比 php 文件优先度高，用户访问 http://www.domain.com 时返回的首页将是 domain.com/index.html 文件，搜索引擎收录的首页也是 index.html 文件。对方却误导你，让你觉得 index.php 文件是首页，页面上有你的链接。这个 index.php 文件只是给你看的。

6．对方根本没链接到你的网站

有的站长检查你的网站有哪些外部链接，然后写邮件给你说，我已经从网站 A 连向你，请你连向我的网站 B，这样是三向链接，比双向链接的效果要好。

我们姑且不论三向链接是否真的比双向链接好。你如果再仔细检查一下链接交换记录，有可能发现网站 A 上的链接其实是你以前和其他网站交换链接时得到的，和现在请求交换链接的这位站长一点关系都没有。他只不过检查了你的外部链接，知道 A 网站上有你的链接，假装那个链接是他做给你的，而有可能你的交换链接比较多，早就忘了是怎么回事了。

7．做一个垃圾网站和你交换链接

有的站长打着三向链接的旗号，要求你的链接必须得连向他的真正的商业性网站，他却从一个垃圾网站链接向你。这种垃圾网站最常见的形式就是垃圾目录，没有什么权重，也没有真实访问流量，是专门用来做友情链接的。

6.9 链接诱饵

随着 SEO 观念和知识的普及，近几年传统外部链接建设手法变得越来越难。链接诱饵是目前比较有效、运用得当时能快速、自然获得链接的方法。

链接诱饵指的是创建有用、有趣、吸引眼球的内容，从而吸引外部链接。从下面几种具体链接诱饵方法可以看到，通过链接诱饵得到的链接最符合好的外部链接标准：全部单向，自愿自发，来源广泛，大部分情况下内容相关度高，有点击流量，链接通常在页面正文中，锚文字自然，等等。

要想吸引链接，就必须吸引其他站长、博主的注意力，让他们觉得链接到你的诱饵页面对自己的用户有很大帮助。普通内容很少能有这种效果，能吸引注意的内容往往需要精心设计、制作，因此才有“诱饵”一词。

6.9.1 链接诱饵的制作

链接诱饵形式五花八门，很难有统一标准和适用于所有情况的模式。最重要的是创意。下面简单讨论几个需要注意的地方。

1．目标对象

毋庸讳言，链接诱饵的最终意义在于链接。而能给予链接的不是普通网民，而是自己拥有网站的站长、博主，以及社会媒体网站上的活跃用户。所以制作链接诱饵时需要研究这些人的需求，而不是普通网民和用户的需求。假设运行的是一个服装销售网站，你很难寄希望于买衣服的普通顾客能带来高质量链接，最多也只是在论坛里夸你两句。能吸引到高质量链接的目标对象很可能是自己也在写博客的服装设计师、造型师，或者其他服装行业的专家，你需要研究这些人会对什么话题感兴趣。

口碑传播、研究用户需求等当然很重要，但不在本书讨论的范围，这里只是讨论链接诱饵问题。

2．刻意与偶然

了解链接诱饵的效果与常用方法后，大部分 SEO 都会尝试制作链接诱饵。不过链接诱饵绝不是百发百中的。花时间、精力制作的自认为不错的诱饵常常没有效果。我自己的博客上就有一些帖子，我自认为挺有创意，应该带来不少链接，结果却完全不然。反

倒是有时候没有带着太强目的性写出的文章效果更好。尽管如此，有目的地制作链接诱饵，如果能达到 20%～30%的成功率就已经不错了。长期坚持下去，积累起来的链接数目将会很惊人。

3．标题写作

再好的内容也需要好的标题首先吸引站长继续看下去。尤其是资源型的链接诱饵，好的标题就是成功的一半。就像在资源型诱饵中所讨论的，“10 大最好的×××”句式就是常见的好标题。

4．去掉广告

链接诱饵的主要目的是吸引站长的注意，不要让站长有任何逆反心理。在可能的情况下去掉诱饵页面上所有广告性质的内容，不要放 Adsense，不要放联署计划代码，也不要推广自己的产品，纯粹以有用、有趣为目标。页面一有商业性，站长们就会对其产出天生的反感，毕竟，站长对各种网站推广手法太熟悉了。

5．易于推荐分享

诱饵页面应该尽量使站长分享方便，比如加上常见的网络书签代码，站长点击一下就可以收录到自己的书签收藏中。工具插件型诱饵要提供现成的 HTML 代码，站长复制一下就能放到自己网站上。

6．设计与排版

页面视觉设计对站长的最初观感有很大影响。一个层次分明、排版整洁的页面，既能让其他站长、博主容易阅读，也吸引他们分享。统计表明，资源型链接诱饵页面上加入图片、视频，大量使用列表都会增加外部链接数量。在这方面，可能吧（www.kenengba.com）是最好的学习对象。

另外，资源型链接诱饵文章长度越长，带来的外部链接也越多。

链接诱饵也是熟能生巧的一门技术，找到自己最擅长的一两种诱饵方法，不断推出同类诱饵，用到极致，长久坚持下去，效果通常比传统外部链接建设方法好得多。

6.9.2 链接诱饵种类和方法

下面讨论几种最常见的吸引链接的诱饵方法。

1．新闻诱饵

行业内有任何新闻，如果你能首先报道，这本身就是链接诱饵。Techcrunch.com 是最典型的英文 IT 新闻网站，在中国 IT 界也同样享有盛名，各种业界新闻八卦，常常在

主流媒体还没反应过来时，就先出现在这个网站上。每一篇新闻出来，都会带来很多链接。

以新闻作为诱饵，必须具备两个特点。一是够快够新，等其他网站都已经报道了，你再添上一笔，就不会有人注意。无论事情大小，第一个报道的总是获得眼球和链接最多的网站。二是够专业化，聚焦于某个垂直领域，切忌贪多。如果大家想看一般性新闻就直接去新浪了，不会到你的网站来看。专业快速的新闻报道最终将使用户产生依赖性，一想到这个行业的新闻，就想到你的网站。

举一个简单的例子，2010 年 4 月份百度推出站长俱乐部。我作为最初几个被邀请的会员之一，得到消息后就在博客上发了一个简单介绍，如图 6-4 所示。注意，本书中查询链接用的 Yahoo!Site Explorer 已于 2011 年 11 月 21 日停止使用。目前还没有合适的替代工具。不过并不影响所显示的链接数量和获得链接的技巧，所以保留 Yahoo!Site Explorer 抓图。

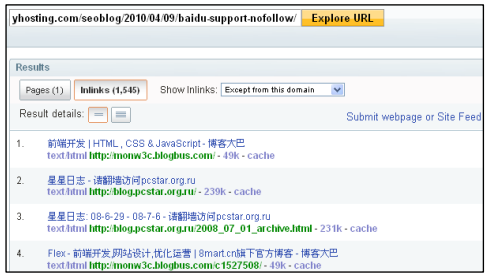


图 6-4　百度站长俱乐部帖子在 Yahoo 显示的外链数

5 月初在雅虎 Site Explorer 查询，已经有 1500 多个外部链接。这可能是网上第一个对百度站长俱乐部的报道。

另一个典型的例子是“谷奥”（原域名是 www.google.org.cn，现已改为 http://www.guao.hk/ ），这是一个专门报道谷歌最新资讯的团队博客，无论是搜索还是 Gmail、谷歌地图，凡是与谷歌有关的新闻，谷奥都快速翻译整理，每天至少都有四五个新帖子。几年前“G 速客”起着同样的作用，不过幻灭离开 G 速客后，其地位迅速下降，谷奥取而代之，如图 6-5 所示。

图 6-5　“谷奥”网站在 Yahoo 显示的外链数

雅虎外链工具显示，谷奥全站有将近 14 万个外部链接。看一下谷奥网站首页就可以知道，友情交换链接并不多，绝大部分外部链接都是靠专业、迅速、勤奋的新闻报道得到的。

2．资源型链接诱饵

这是最简单也最有效的一类链接诱饵。提供某一个话题的全面、深入资源，就能成为吸引外部链接的强大工具。所谓“资源”，既可以是一篇深入探讨的教程或文章，也可以是连向其他资源的列表。

比如著名的 SEOBook.com 博主 Aaron Wall 写的“博客 SEO 指南”，获得了 2000 多个外部链接。这是一篇非常深入、全面地探讨博客 SEO 方法的文章，篇幅很长，主题集中，在 SEO 领域有广泛影响，如图 6-6 所示。

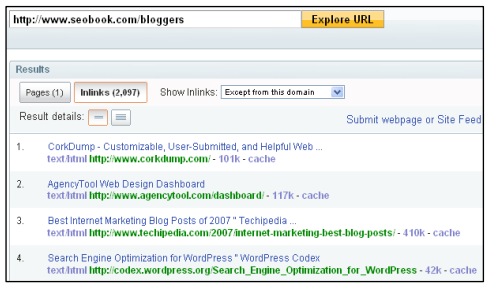


图 6-6 “博客 SEO 指南”英文原文外链数

我得到 Aaron Wall 的授权，将这篇指南翻译成中文，发表在博客上，也获得了 100 多个外部链接，其中包括来自 SEOBook.com 的指南原页面。网上没有保留原出处的转载、抄袭就更多了，如图 6-7 所示。

图 6-7 “博客 SEO 指南”中文翻译外链数

“可能吧”（www.kenengba.com）是一个非常善于使用资源型链接诱饵的博客。博主 Jason 写了很多篇幅长、有深度，同时关注热门话题的帖子。而且“可能吧”的所有文章中排版展现突出，大小标题清晰严谨，大量使用插图，以颜色区分标题与正文，使用红色等鲜明颜色突出正文中的重点语句。这样的排版使访问的人一看到就觉得赏心悦目，内容又非常深入，对相关话题做了很完整的分析，自然吸引了大量关注，如图 6-8 所示。

图 6-8 “可能吧”页面排版出色

以“网络文化背后的法则”帖子为例，雅虎工具显示 80 多个外部链接，如图 6-9 所示。

图 6-9 “网络文化背后的法则”帖子外链数

另外一篇“微博客里的信息干扰”，90 多个外部链接，如图 6-10 所示。

图 6-10 “微博客里的信息干扰”帖子外链数

资源型诱饵也可以是简单的资源列表形式，月光博客（www.williamlong.info）就深

谙此道。读者留心的话会发现，月光博客上经常看到 10 大这个，15 个那个这种资源列表文章。这些列表性质的文章非常受欢迎。例如“16 个扩大博客影响力的方法”这篇有 44 个外部链接，如图 6-11 所示。

图 6-11 “16 个扩大博客影响力的方法”帖子外链数

“提高浏览体验的 50 个最佳 FireFox 扩展插件”这篇有近 300 个外部链接，如图 6-12 所示。

“可能吧”也有不少这种列表资源，图 6-13 是“可能吧”权重比较高的内页的一部分。读者可以看到列表类资源占到一半。

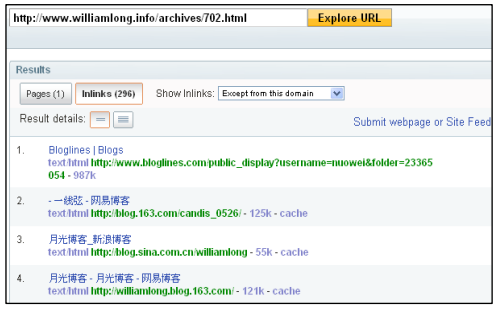


图 6-12 “提高浏览体验的 50 个最佳 FireFox 扩展插件”帖子外链数

图 6-13 “可能吧”权重帖子较高

再比如 SEObbs.net 网站上的 SEO 工具列表页面，乐思蜀把收集到的 SEO 工具列出名称，简单说明，链到相应工具页面，有 80 个外部链接，如图 6-14 所示。

SEO 人员都知道吸引外部链接很难。其实，制作这种资源列表难吗？只要用心收集整理，一点都不难。

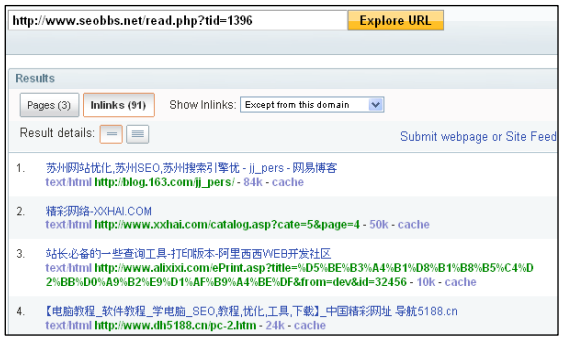


图 6-14　SEObbs.net 工具列表页面外链数

3．争议性话题

带有争议性的话题显然能吸引到眼球，而且经常能吸引到争议双方你来我往进行辩论，围观者传播、评论。

2007 年，著名网上创业家 Jason Calacanis 发表了一篇帖子，指 SEO 是胡说八道，立即在英文 SEO 行业掀起争论热潮。很多成名 SEO 纷纷指出 Jason Calacanis 言论的逻辑错误，实际上 Jason Calacanis 自己公司网站上把 SEO 列为公司优势之一，可见 Jason Calacanis 并不真的认为 SEO 是胡说八道，只不过是想为他新推出的网站造势、造链接而已。一部分 SEO 人员明白这一点，在讨论 Jason Calacanis 的文章时，刻意不给链接。不过还是有很多辩论帖子指向了 Jason Calacanis 个人博客及公司网站。

给行业权威、名人挑毛病，甚至大骂一顿，是常见的能带来外部链接的争议话题。2010 年 5 月初，著名博客麦田发表一篇题为“警惕韩寒”的帖子，认为韩寒的文章是在迎合大众情绪。这里无意讨论是否需要警惕韩寒，麦田和韩寒都是我非常喜欢的博客。麦田本人不是 SEO 人员，也不会心里想着链接才写这篇帖子，只是无意间成为很好的链接诱饵例子。

麦田的这篇帖子发表后引起的争议不小，带来的链接也不少。5 月中旬，帖子发出一个星期，通过雅虎外链工具查询，仅麦田在新浪的这篇帖子就得到了近 200 个外部链接，如图 6-15 所示。

图 6-15　麦田在新浪“警惕韩寒”帖子外链数

当然，作为链接诱饵的争议性话题，还是得在理性的基础上讨论，千万不要捕风捉影，无端谩骂。

有时候争议性话题也不一定这么严肃。笔者在 2010 年 1 月发过一篇博客：修改页面标题是否影响排名。因为国内 SEO 行业有这么一种说法，修改页面标题就会被搜索引擎降权，使排名下降。笔者对此有怀疑，所以做了个实验，然后发帖说明实验结果。

修改标题是否影响排名，但现在为止还是个争议性话题。我的实验只能说明一部分网站不受影响，比如我的博客这种有一定权重的域名。对新站、小站是否有影响还是有不同意见。这篇有些争议性的帖子也得到了 60 个外部链接，如图 6-16 所示。

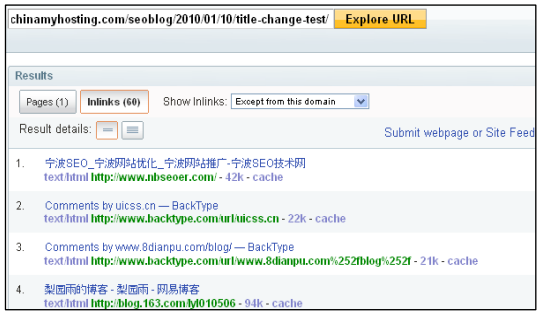


图 6-16 “修改页面标题是否影响排名”帖子外链数

4．线上工具

这是 SEO 最熟悉的链接诱饵，几乎每天要用。网上 SEO 工具种类繁多，诸如查询 PR 值、查询收录数、计算页面关键词密度、查询百度指数、查询相关关键词等，既有搜索引擎提供的工具，也有站长们自行开发的工具。本书 SEO 工具部分介绍的大多数都是线上工具，每一个工具都有很强大的外部链接。原因很简单，这些工具是 SEO 人员所需要的，站长们会在自己的博客上、网站上、论坛里推荐给其他人。

举两个中文 SEO 工具。身在加拿大的 David Yin 的网站优化推广博客（http://seo.g2soft.net/）很有参考价值。他开发的页面重定向检查工具有 50 个外部链接，如图 6-17 所示。

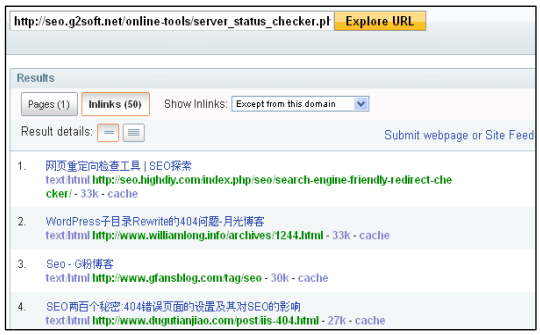


图 6-17 David Yin 重定向检查工具外链数

读者从图 6-17 中可以看到，链接过来的不乏 SEO 及 IT 行业的有名博客，权重都不错。

另一个最近发现的工具是莫大（meta.cn）。这个工具非常新颖，很难用一两句话说清楚，建议大家直接去网站试用一下。大家可以看到 meta.cn 网站完全没有友情链接，但是却有 14 000 多个外部链接，其中包括笔者本人的博客，如图 6-18 所示。

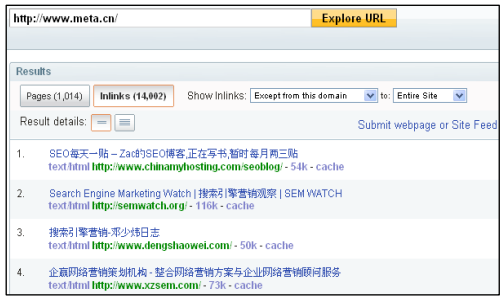


图 6-18　meta.cn 外链数

当然，线上工具不仅限于 SEO 工具。如图 6-19 所示是一个估算博客价值的线上小工具。其原理是计算博客在 Technorati（一个曾经很流行的博客搜索服务）的外部链接，然后比照博客 Weblogs 被 AOL 收购的价格和外部链接，计算出其他博客的潜在价值。

工具地址是：

http://www.business-opportunities.biz/projects/how-much-is-your-blog-worth/

图 6-19　博客估值线上工具

显然这个线上工具没有什么实际价值，只是好玩而已。就算你的博客被估值几万美金，几乎可以肯定没有人会用几万美金买你的博客。不过看一下自己博客的潜在价值，毕竟是一件挺有意思的事，所以很多人推荐使用。这个工具页面有惊人的 100 多万外部链接，如图 6-20 所示。

线上工具也不一定要和网站有什么关系，其他领域同样可以制作出有用或者好玩的线上工具，比如计算贷款利息，计算一个人是否超重，计算一个婴儿从出生长到六岁一共要喝多少罐奶粉，计算预产期等。关键在于开阔思路，线上工具有太多可能性。

5．插件

对有技术基础的公司和 SEO 来说，写插件也是一个非常有效的链接诱饵。

最简单的插件可以是给博客、购物车、论坛等开发功能性插件。这方面最好的例子是擅长写 WordPress 插件的 Yoast。这个页面是 Yoast 给 WordPress 开发的面包屑导航插件：

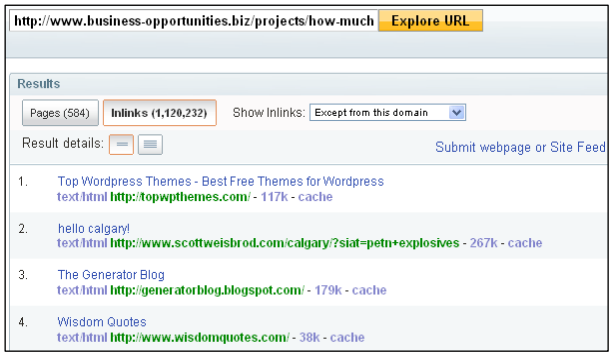


图 6-20　博客估值线上工具外链数

http://yoast.com/wordpress/breadcrumbs/

安装这个插件后，博客会自动生成面包屑导航。这个插件发布页面有 3000 多个外部链接，如图 6-21 所示。

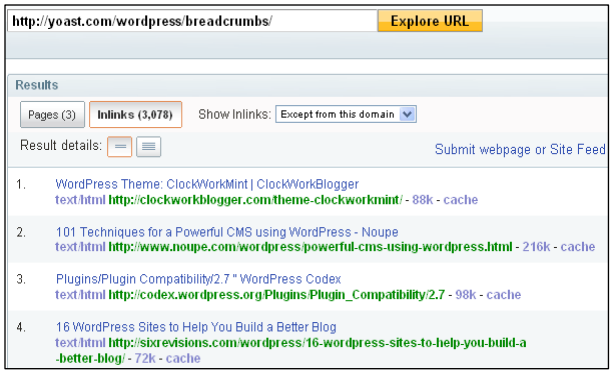


图 6-21　面包屑导航插件页面外链数

另一个更强大的链接诱饵是 Yoast 给 Cononical 标签开发的功能插件。Google 等搜索引擎推出 Cononical 标签后，Yoast 几天之内就推出了这个插件，给 WordPress 博客上大量重复页面，如按时间存档、分类存档等自动加上 Cononical 标签。这个插件页面有 600 多个外部链接，如图 6-22 所示。

最让人惊叹的是，插件竟然吸引了来自 Google、雅虎官方博客的链接，也有 Matt Cutts 的链接，还有来自纽约时报、Youtube 的链接，最著名的 SEO 服务商 SEO Moz 的链接，SEO 行业第一人丹尼苏利文主持的 searchengineland.com 的链接等。这些网站权重都非常高。Yoast 写的这个插件给自己域名带来的权重提升是非常惊人的。对 Yoast 来说，开发这样一个插件不会费太多事，重要的是敏感的嗅觉。

图 6-22　Cononical 插件页面外链数

再举一个国内 WordPress 专家写的插件。Google 推出 NoFollow 标签后，大部分博客评论链接都自动加上了 NF 标签，不过百度是否支持 NF 在过去几年都没有定论。为了确保博客留言中的垃圾链接不能传递权重，仅使用 NF 可能不够，最好能通过程序转向完全断绝权重传递。

笔者在线上与中文 WP 专家水煮鱼聊了几句后，几分钟之内水煮鱼就写好了这个插件。真是难者不会，会者不难。水煮鱼发布这个插件的页面现在有 70 多个外部链接，如图 6-23 所示。

还有一类小工具作为链接诱饵更加强大。工具开发者提供一段 HTML 代码，站长把这段代码放在自己的网站上，就可以实现某种特殊功能。比如前几年颇为流行的 MyBlogLog，博主把 MyBlogLog 提供的代码放在博客侧栏中，这段代码就显示出有哪些人访问了这个博客。当然访问的也必须是 MyBlogLog 的注册会员才能显示。如图 6-24 所示，这个功能区块右下角会自动生成一个链接，指向 MyBlogLog 网站。

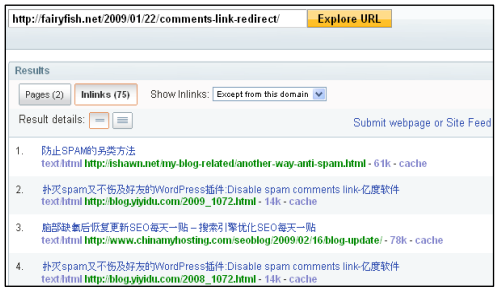


图 6-23　WP 评论转向插件外链数

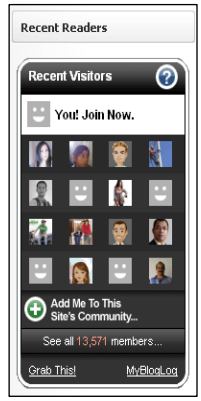


图 6-24　MyBlogLog 插件

使用 MyBlogLog 的博客越多，指向 MyBlogLog 的外部链接也就越多。对博主来说，MyBlogLog 提供了一个很有意思的功能，也不会介意一个导出链接。

上面提到的博客估值工具也有这个功能。博主除了在线上版工具看一下自己博客估值外，还可以把工具生成的代码放在自己网站上，代码自动显示本博客值多少钱，如图 6-25 所示。

当然估值下面还有一个链接指向原工具页面。抓图中“How much is your blog worth?”这句话就是链接，如图 6-26 所示。

图 6-25　博客估值工具生成的 HTML 代码

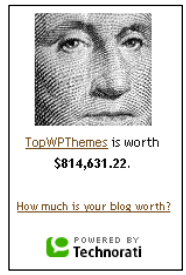


图 6-26　博客估值插件显示的图标及链接

很多博主为了好玩或稍微炫耀一下，会把估值代码放在自己博客上。这个工具页面的 100 多万链接，没有这种插入代码式的工具恐怕也做不到。

这个博客估值工具也有中文替代品，如图 6-27 所示。

同样，用户单击“估值是”按钮后，工具除了显示网站价值，还给了一段代码，站长可以把代码放在自己的网站上，炫耀一下自己网站的价值，如图 6-28 所示。

图 6-27　中文博客估值工具线上版

图 6-28　中文博客估值工具生成的 HTML 代码

从代码里我们就可以看到，“我的网站价值是”这几个字被做成了链接，指向工具

页面。用雅虎外链工具查询一下可以看到，工具页面有 4000 多个外部链接，如图 6-29 所示。

图 6-29　中文博客估值工具外链数

6．利益吸引

除了内容有用，站长提供链接就能得到某种好处，也是形成链接诱饵的不错方法。这里所说的好处不是金钱，而是其他方面的利益。

几个国内 SEO 网站免费给其他站长提供 SEO 诊断和咨询，唯一的条件是接受诊断的网站需要提供一个单向链接，起过一些效果。

投票、排名、比赛也是常见的利益吸引方法。如 SEOMoz 举行的 Web 2.0 大奖，入围各奖项的网站为了给自己拉票，都会在各自网站上宣传这项竞赛，放上链接方便用户到大奖赛页面投票。最后获奖的网站，也免不了链接到奖项官方网站炫耀一下自己获奖了。当然在评比过程中也引起了并没有参与奖项的网站和媒体的关注，给予了更多链接。SEOMoz 的 Web 2.0 大奖页面现在有 23 000 多个链接，如图 6-30 所示。

图 6-30　SEOMoz Web2.0 大奖页面外链数

另一个例子是在中文 SEO 行业影响深远的第一届 SEO 大赛——渡虎谷大赛。除了参赛网站需要链接到大赛官方网站以外，站长社区的谈论、各参赛网站优劣评论都吸引了很多外部链接。现在渡虎谷大赛官方网站已经下线，但是雅虎外链工具还显示有 400 多个外部链接，如图 6-31 所示。

图 6-31　渡虎谷大赛官方网站外链数

7．幽默搞笑

笑话和段子是网上传播最快的内容之一。幽默搞笑的内容也经常吸引到很多外部链接。一个很好的例子是这个网站：

www.worldsgreatestbusinessmind.com/

用户在网站上输入自己的域名，网站将实时生成显示一段电视新闻报道片，其中有很多镜头是根据用户输入的域名实时修改的，所以效果是把用户域名嵌入到新闻片段里的电视屏幕、报纸、广告牌等地方，如同一段真实的宣传报道网站的新闻片段。如图 6-32 所示是输入点石域名后生成视频的一个镜头。

图 6-32　搞笑广告生成网站

这个网站在某种意义上说也可以称为一个工具，不过是没有工具价值的工具，只是好玩，所以笔者把它放在幽默搞笑类的链接诱饵中。站长把自己的域名嵌入视频，自然会把视频 URL 放在自己的网站上，告诉其他人，博用户一笑。雅虎外链工具显示，这个搞笑工具网站有 3200 多个外部链接，如图 6-33 所示。

幽默搞笑也不一定要这么复杂。每年愚人节很多博客都会愚一把，笔者也如此。如图 6-34 所示是笔者博客上 2009 年愚人节帖子的外部链接，虽然只有 16 个，不过得来非常简单，只不过是一个愚人节笑话而已。

著名网赚博客 Shoemoney 曾经在他的博客上发过自己以前很胖时候的照片，告诉大家自己以前是个胖子，如图 6-35 所示。

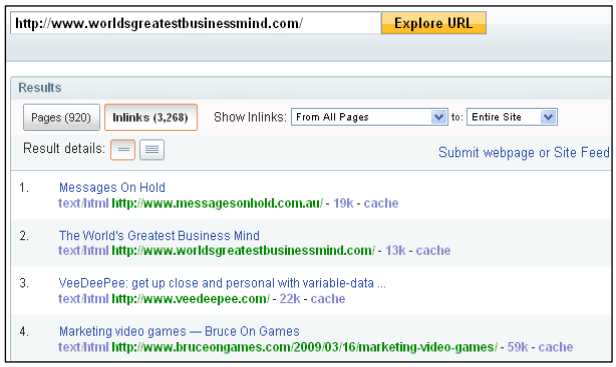


图 6-33　搞笑广告制作网站外链数

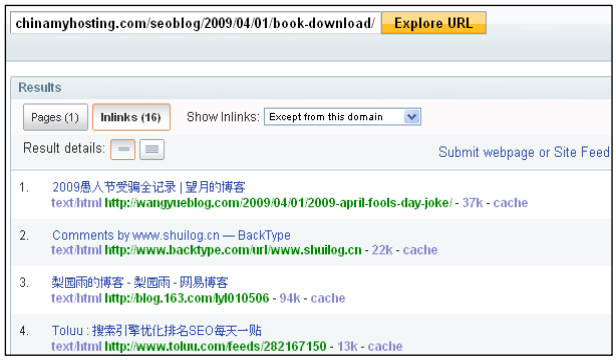


图 6-34　愚人节帖子外链数

图 6-35　Shoemoney 讲述自己以前很胖的帖子

这个帖子与 SEO、网赚没什么大关系，但是也很好玩，吸引了 100 多个外部链接，其中有不少权重很高的著名博客，如图 6-36 所示。

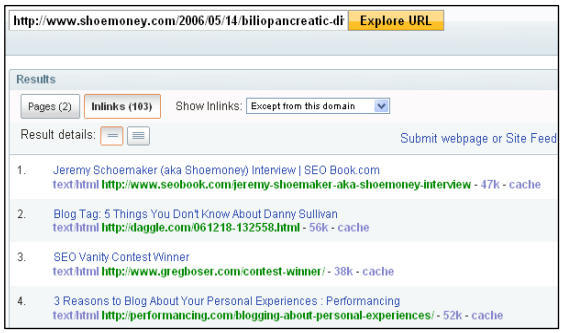


图 6-36　Shoemoney 帖子外链数

6.9.3　链接诱饵之度

虽然链接诱饵吸引来的链接都是自愿、单向的，但也有个“度”的问题。做得过火，搜索引擎可能会认为有意图操纵排名之嫌。举个比较有名的例子，读者可以大致了解搜索引擎的底线在哪里。

很有才华的前 SEOMoz 员工 Matthew Inman 自己建了一个网上交友网站 Mingle2，并想了一个很好的诱饵吸引外部链接。他编写了一些线上测试和问答程序，用户做完题目后，程序显示成绩，并且提供一段 HTML 代码，用户可以把这段代码放在自己的博客上，显示出测试分数。当然，代码中包含指向 Mingle2 的链接，锚文字是免费网上交友（free online dating）、网上交友（online dating）等。这是典型的小工具链接诱饵。

Matthew 写的测试题都挺有意思，比如你能在月球上生存多久？你能在真空中生存多久？测试一下你在博客里是不是话太多了？如果被困在家里，你能生存多久？你在 5 分钟内可以想出多少个国家名称等。据 Matthew 说，他写了几十个这样的测试题，获得很多外部链接。

他的个人网站取得了很大成功，网站被另外一个网上交友网站 JustSayHi 收购，人也被挖了过去。Matthew 在新公司继续发挥自己工具诱饵的特长，把 JustSayHi 网站排名迅速提高了。

到这里都挺好，然后麻烦开始了。

JustSayHi 的母公司 Next Internet 又收购了一些其他网站，包括贷款、卖药等竞争激烈、黑帽充斥的行业。很自然，母公司希望 Matthew 使用同样的方法帮助推广新收购的网站。Matthew 就在自己的工具代码中放上了那些新网站的链接。没过多久，不但那些新网站排名没上去，连原来已经排名第一的网站 JustSayHi 也被惩罚了，连搜索网站名称时都没有排在前面。

好在 Matthew 与 Google 内部人员有些联系，经过询问后，知道被惩罚的原因主要是方法用过了头，工具插件中的链接不仅指向原来开发诱饵的 JustSayHi 网站，还开始指向第三方网站，有付费链接的嫌疑。链接锚文字也有关键词堆积嫌疑。

Matthew 提交了重新收录请求，承认了自己的错误，不过并没有使网站从惩罚中逃出来。

从这个案例可以看出，有创意的链接诱饵也有个界限。使用得当，SEO 效果威力惊人。像 Matthew 的工具诱饵，几十天内造出几十万外部链接。但使用过头，搜索引擎很生气，后果很严重。

从 Matthew 自己的介绍和 Matt Cutts 对这个案例的非直接评论看，Google（相信其他搜索引擎也类似）对链接诱饵（尤其是工具型诱饵）有些质量要求：

- 链接是否是隐藏的？工具程序生成的代码把链接设置为隐藏文字或链接，或者把链接放在 NoScript 之类的代码中。
- 相关性有多高？一个好的相关性高的工具例子是，Ubuntu 小工具显示 Ubuntu（一种基于 Linux 的操作系统）离正式发布日期还剩下多少天，就是一个简单的倒数计时器，链接向 Ubuntu.com。
- 链接是连到开发工具的那个网站，还是链接到一个完全无关的第三方网站？链接向第三方很可能意味着这是一个出售或付费的链接。指向第三方的链接，很容易被检测到。
- 链接文字是否过度优化？链接锚文字是工具开发网站的名称很正常，堆积上关键词就开始可疑了。
- 工具中包含有多少个链接？如果有一堆链接，对工具功能有什么帮助？还是仅仅为了制造链接？
- 放上工具代码的站长是否知道工具中含有链接？有的工具说明网页把链接的事藏在用户条款的第 26 条等之类的地方，实际上用户不会看到用户条款，可能就不知道有这样一个链接。嵌入工具或插件的人必须知道代码中有链接，知道链接到什么地方，所以链接是带有编辑投票意义的。

6.10 其他常规外链建设方法

本节简单介绍其他常规外链技术方法。

1. 自己网站

新网站建设外部链接比较困难，在没有 Google PR 值、收录也不充分的情况下，就连最简单的友情链接交换也不容易。大部分站长交换链接时还要检查网站 PR 值，首页 PR 为零，很难说服对方与自己交换。

我个人的方法是新站先在自己的老网站上做几个链接，过几个月后 PR 值提升到一定程度，收录也比较正常后再开始交换链接。拥有一两个至少中等 PR 值的网站，或者养一些信息类网站、博客等，对新网站建设很有作用。

2．博客

这里所说的博客是指建在企业网站独立域名上的博客，而不是第三方免费博客平台。

现在站长都知道外部链接的重要性，一般商业性网站、企业网站很少自发链接到其他网站，比较愿意给链接的只剩下几种网站，博客是其中最容易的一种。至少到目前为止，相关博客之间互相引用、评论还比较常见。博客通常是比较个人化的内容，商业性较低，积极参与本行业博客圈，与权威博主讨论话题还是能够吸引到不少高质量的外部链接。

建议把博客放在域名的二级目录上，而不要放在子域名上，因为通过博客获得的外部链接，有可能连到博客首页，也有可能是博客帖子页，如果博客放在子域名上，这些外部链接的权重将只传递到子域名而不是主域名。

3．文章发表

网上有很多自由发表文章的地方，如第三方免费博客平台，各大门户网站及区域地方性门户都提供博客平台。有的网站也提供普通文章发表账号。每个行业几乎都会有 SNS 类网站，用户可以免费注册，发表日志、话题。

还有一类专门收集文章的网站，用户可以建立账号，自由发表文章，其他站长也会到这种文章目录网站寻找好的内容转载到自己网站上。有很多这种英文文章目录，搜索一下“article directory”之类的词就能找到几百个。

上面这些可以发表文章的地方，SEO 都可以建立和积累账号，自己有博客帖子或窍门、信息类文章时，能够在众多网站发表。每篇文章中都可以留下指向主网站的链接，既可以是版权声明中指向网站首页的链接，也可以是文章正文中相关内容处指向内页的链接。

有的文章发表网站对包含的链接比较敏感，可能会在审核时删除链接。不过只要做得不过分，每篇文章中只含有少量链接，而且只在必要时才留下链接，坚持提供有意义、有价值的文章，审核通过率会保持在一定水平。

文章发表是自己可以在一定程度上控制，而且有可能做到很大数量外部链接的方法之一。要注意的是，文章首先要在自己的网站上发表，等被搜索引擎收录后再到其他地方发表。最好有链接连到自己网站上的原出处，帮助搜索引擎判断哪个是原创版本。如果有时间和精力，可以将自己网站上的原始版本稍作修改后在其他网站发表，尽量避免成为复制内容，从而提高收录可能性。

4．论坛博客留言

在论坛留言中包含外部链接也是常见的方法，很多论坛签名也允许加上链接。站长和 SEO 行业论坛经常不允许在正文中留链接，甚至签名也不允许链接，比如点石论坛、

英文的站长世界（webmasterworld.com）。站长世界更为极端，连具体网站名称、非链接形式的 URL 都不能出现。

但是网上还有很多其他行业论坛管理员并没有像 SEO 一样对链接这么敏感，还是允许留下链接的，而且很多论坛中的链接没有加 NF 属性。当然，在论坛中留链接，并不是鼓励大家到论坛里发垃圾留言，首先还是要参与讨论，提出有建设性的观点。

博客评论中的链接现在很少见到没有 NF 属性的，但博客留言是引起博主注意、与其他读者沟通联系的很好方法。只要在博客留言中言之有物，也能在圈子里留下美名，引起注意，间接吸引其他人访问你的网站。

现在有一部分博主对 NF 属性的大量使用颇为反感，认为 NF 属性破坏了博客互动的本质，所以在一小部分博主中兴起了 DoFollow 活动，使用插件将博客软件中默认设置的 NF 标签去掉。可以在 http://followlist.com/寻找 DoFollow 博客，在这种博客上的留言直接链接效果更好一些。

长远来说，可以通过有价值的评论引起博主注意，潜在外部链接价值比任何形式的 DoFollow 留言高得多。笔者的博客帖子经常会链接到其他人的博客和网站，很多人引起我的注意，订阅他们的博客，就是通过留言。

5．合作伙伴

所有公司企业都有各种各样的合作伙伴，供货商、客户、投资方、原材料供应商、生意伙伴都可以是外部链接来源。企业也可能参加了行业组织，隶属于某个地区性联盟，也可能与某些政府部门有特殊合作关系，与学校科研机构有学术合作关系，等等。

每一个合作伙伴都是可以请求和获得外部链接的来源，而且来自合作伙伴的链接通常质量比较高。在谈合作时简单的一句话，就可能带来一个高质量的外部链接。

个人站长也一定有亲戚朋友，有同属于一个圈子的其他站长朋友。参加行业聚会、研讨会时，多与其他站长做线下沟通与联系。很多时候通过邮件、QQ 都搞不定的链接，见个面聊几句，一起吃顿饭就全部搞定了。这种线下的日常合作伙伴往往是网站建设外部链接时所忽略的，但其实很容易获得的外部链接来源。

6．网摘书签

网上有很多线上书签服务，如最著名的 delicious.com、QQ 书签、百度搜藏、Google 书签等。站长可以注册这些线上书签服务，网站上有新内容时，先存到自己的书签账号中。网站上的内容页还可以加上主要书签网站代码，方便其他用户收藏书签。被收藏的书签页面会出现在书签网站上，形成外部链接。书签网站本身还有很多用户，能够通过热门类别、分类、tag 集合等方式发现你的内容。

目前很多书签网站都在链接中加了 NF 属性，但也有一些没有加。有的我们都以为加了 NF 属性的书签网站，其实在某些特定页面上还保留了没有 NF 的链接。SEO 人员要多积累，善于发现。

而且就算书签中的网址加了 NF，也可能带来点击流量。用户通过书签来到你的网站，如果觉得内容质量确实高，还可能间接带来外部链接。

7．百科及问答类网站

现在百科和问答类服务很热门，常见的包括百度百科、百度知道、天涯问答、雅虎问答等。这种线上服务页面中通常允许相关链接，方便用户查看更多资料。如果你的网站有与词条相关性很高的资源内容，可以加在外部链接、扩展阅读部分，如图 6-37 所示。

图 6-37　百度百科词条中的外部链接

为了预防垃圾链接，这种服务对条目的编辑，尤其是增加链接都很敏感，所以使用时一定要实事求是，只有在你的页面确实对条目是很好的补充说明时再进行编辑。很多站长的经验表明，有一定权威度的网站、确实有价值的内容，在诸如百度百科中审核通过的机会并不低。

同样的方法也适用于维基类网站。

8．客座及团队博客

近一两年客座和团队博客多了起来。几年前点石博客创建时还很少看见其他团队博客，一般博主都是单打独斗的。最近国内互联网行业出现了几个非常不错的团队博客，或者接受客座帖子的博客：

- SEMWatch.org，由几个成员组成，同时接受投稿。
- “可能吧”除了最初的博主 Jason 外，现在也经常出现其他团队成员的博客。
- 著名的 IT 博客“月光博客”，最近也接受客座投稿，而且数量不小。
- 谷奥博客是另一个人气非常高的团队博客，同时接受投稿或用户推荐内容。

这些接受投稿的博客，都会在发表客座帖子时链接到作者网站，并且保留帖子正文中的链接。团队博客也往往以作者落款的形式，允许链接到作者自己的主博客。这其实

是一个非常好的外部链接机会。熟悉 IT 和互联网行业的人都知道，月光博客权重高、影响大，要想自己的网站链接出现在月光博客上，难度很大。与月光博客交换链接也很困难，实际上月光好像不交换链接。以客座帖子的形式，使自己的文章和链接出现在月光博客上的价值不言而喻。

每个行业都会有几个权威博客。也许他们还不接受客座形式，SEO 人员完全可以尝试联系博主，建议以客座形式提供有价值的原创内容，对方不一定就会拒绝。这种合作也许还能在你的行业内掀起一股创新热潮。

9．检查竞争对手网站

竞争对手网站链接来源是个取之不尽的宝藏。强烈建议 SEO 人员用雅虎外链工具或者后面 SEO 工具章节介绍的 SEOMoz 和 Majestic SEO 检查主要竞争对手外链情况，看看竞争对手有哪些外链？为什么人家会给予外链？是交换还是链接诱饵？还是新闻或者购买的链接？你能不能说服这些网站也给你一个链接？既然他们能给竞争对手链接，你能成功获得链接的可能性也不会很小。

强有力的竞争对手可能有成千上万的外部链接。上面提到的几个工具，都是大致按重要程度列出外部链接。SEO 人员没有其他任务时，把竞争对手外链过一遍，一定能发现好的来源。如果你能把竞争对手具备的外链也都得到，再加上从其他渠道获得的竞争对手没有的外部链接，你的排名和流量超过竞争对手就是自然而然的事情了。

10．网络广告

网络广告早在搜索引擎出现之前就已经存在了，以后还会继续存在下去。只要没有使用 JS 脚本或通过程序转向，普通 HTML 链接做成的广告就是正常的外部链接。

虽然网络广告经常是旗帜、按钮等图片形式，一般认为链接效果不如文字链接，但图片链接也还是链接，图片 ALT 文字大致相当于锚文字，网络广告使用得当，除了能带来点击流量、品牌价值，还能带来外部链接，何乐而不为？

网上有很多网站，尤其是信息类和个人网站，有不错的内容和流量，却整天发愁怎么赢利。主动联系这些没有清晰赢利模式的网站投放广告，有时能拿到非常便宜的价格，几十、几百块一个月，也许就能在一个流量不错的论坛或信息网站获得全站链接。要找到这种网站也相对容易，只要搜索相关关键词，那些没有出售自己的产品，大量出现 Google Adsense、各种联盟广告的网站，都是不错的备选广告目标。

11．购买链接

如果网络广告以文字链接的形式出现，就变成购买链接，这是一个从 SEO 角度看比较灰色的模糊地带。搜索引擎鲜明反对以排名为目的的购买链接，检测出链接是购买链接的，会把链接的投票权重取消，还可能对买卖双方网站实施惩罚。

但是检测链接买卖是一件非常困难的事情，而且链接买卖也经常混杂着多重目的，不一定仅是为了影响排名。主要以点击流量为目的，又能带来曝光度、外部链接，这样的购买链接也不妨一试。

要注意的是，尽量不要在明显的链接交易平台或论坛找卖家。这种大量卖家和买家聚集的地方，通常都是以 SEO 为目的的。在那里卖链接的网站也不会仅仅卖一两个链接，而是能卖多少就卖多少，不管质量。这种链接买卖网络可以比较容易地被检测出来。搜索引擎通过人工方式，确定少数几个存在链接买卖的网站，就能通过链接关系挖出一大堆参与其中的网站。想要尝试购买链接，最好搜索相关关键词，直接联系排名不错的网站站主，对方没有主动出售链接，效果则更好。

12．社会化媒体网站

除了线上书签网摘服务以外，其他社会媒体网站也有助于外部链接建设，包括内容分享、社交网络等形式的网站。

首先，社会化媒体网站上引用、推荐、评论网址很常见，在诸如新浪、腾讯微博、Twitter、Facebook 网站上，一个有很多听众的账号贴出网址，经常会带来大量点击访问。如果你的页面内容吸引了这些访问者，引起他们的评论、反馈欲望，就可能间接带来外部链接。

其次，虽然大部分社会化网络上出现的导出链接都加了 NF 属性，但是搜索引擎到底怎样看待这种链接还是未知数，并不一定像公开谈论的那样完全忽略 NF 链接。社会化网站上大量链接及用户点击的流动，很可能是最新的排名信号之一。比较明显的是，很多人注意到维基百科（wikipedia.org）上的外部链接虽然一律加了 NF 属性，但似乎对页面排名还是有帮助，其原因既可能是搜索引擎对维基百科上的带有 NF 标签的链接做了特殊处理，也可能是点击流量间接带来其他地方的外部链接。

有的社会化媒体网站还允许用户建立不带 NF 属性的链接，比如用户信息页面（Profile）通常允许加上用户的个人网站地址。如 www.scribd.com 是一个文档分享网站，它的用户信息页面链接就是正常的 Follow 链接，如图 6-38 所示。

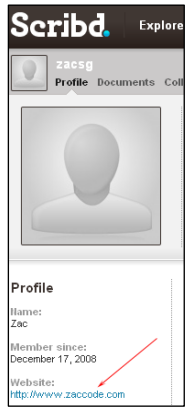


图 6-38　scribd.com 用户信息页面中的链接

13．联署计划

联署计划也就是站长们熟悉的网站联盟。广告商给参与计划的站长提供一个跟踪代码，站长把这段跟踪代码放在自己的网站上。用户点击链接来到广告商网站产生购买或其他转化，广告商支付佣金给站长。

在某些情况下，联署计划跟踪链接也能形成外部链接。只要联署计划程序运行于广告商自己的域名上，并且跟踪链接做 301 转向到首页或相应的产品页面，这些联署计划

链接就变成了普通外部链接。如果使用第三方网站联盟服务，联署计划链接实际上是指向第三方服务商域名，再转向到广告商网站，在第三方服务商网站上通常还要做 cookie 设置、跟踪脚本转向等，使得链接权重不能顺利传递，就不能形成外部链接。

14．CMS 系统模板

网上有大量免费开源 CMS 系统，很多网站尤其是个人网站和小企业建站，都是使用现成的 CMS 系统，如 dede，shopex，WordPress，ecshop，OSCommerce，zen-cart 等。

使用这些 CMS 系统的网站非常多，模板需求量自然很大。熟悉特定 CMS 系统模板制作的站长可以设计不同风格的模板，供其他站长免费使用。唯一的条件是，在模板版权声明处留下设计者链接，如图 6-39 所示。

图 6-39　点石博客页脚处模板设计者链接

这种方法是外部链接建设非常有效的方式，往往能带来大量链接，因为模板上的版权声明通常会出现在网站所有页面。

我在博客中曾经提到过崔凯博客这个案例。在研究崔凯博客的外部链接时，我就惊讶地发现他的博客有十多万个外部链接，绝大部分来自于他设计的 WP 模板，主要是英文模板，如图 6-40 所示。

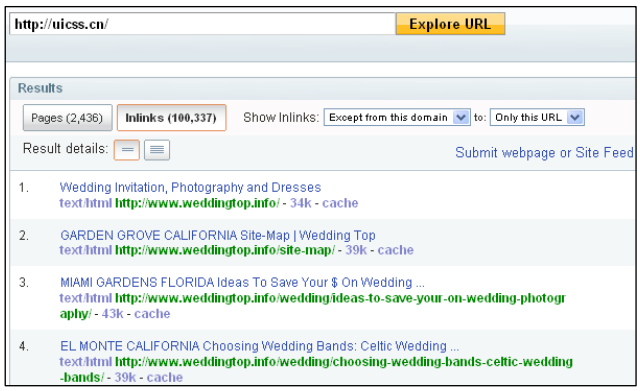


图 6-40　崔凯博客外链数

到 WP 官方网站上看看就会知道，很多模板下载次数非常大，被使用的也很广泛。就算有一部分使用者删除了模板中的设计者链接，保留链接的数量还是很庞大。点石博客模板设计网站的链接数如图 6-41 所示。

惊人的六百多万！这要是靠交换链接得交换到什么时候？模板制作类似于链接诱饵，或者也可以理解为链接诱饵的一种。

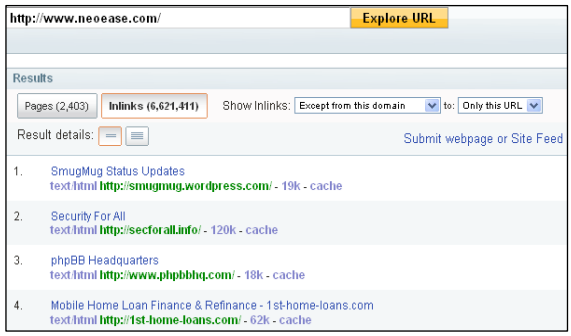


图 6-41　点石博客模板设计者网站外链数

需要注意的是，模板链接通常是全站链接，这是搜索引擎比较敏感的链接种类。所以建议锚文字不要商业化，用网站正式名称或设计者网名就可以了，不要使用要优化的目标关键词，尤其不要期望通过模板中的锚文字优化诸如伟哥、仿品、六合彩这种关键词。重要的是提高权重，扩展链接来源总域名数。

15．新闻稿发布

网上有一些专门的新闻稿发布网站，以英文居多。在 Google 搜索一下“press release”，就能找到很多新闻发布网站。

这些新闻发布网站允许用户注册，提交和发布新闻稿件。新闻结尾处通常允许留下发布者网站链接，正文中还可能允许留下一两个链接。新闻稿除了会出现在新闻发布网站本身，新闻发布服务商还会把新闻稿推到与自己有渠道合作关系的其他新闻类网站上。有些小的新闻网站会自动去新闻发布网站上抓取内容放在自己网站上。

有的新闻发布需要付费，比较有名的如 PRleap.com、PRNewsWire.com 等。支付一定费用后，新闻发布服务将保证稿件出现在一定数量的渠道网站上。也有很多免费新闻发布，当然他们的发布渠道很少，不过最差的情况也至少会出现在新闻发布服务自己的网站上。

新闻网站通常权重比较高，所以是比较好的外部链接。就算是付费服务，也只是一次性费用，就能永久保留链接。这比某些购买链接划算，通常买链接是按月付费的。

近几年中文新闻发布服务也出现了一些，比如 prexpress.cn。

16．媒体公关

有公关实力和渠道的企业网站，在进行公关活动时不要忘了，除了提高曝光度，公关也可能成为外部链接建设的强有力方法。在发布新闻时提到公司网址，应该是公关标准流程之一。很多时候公司高层的一举一动都是新闻媒体追逐的对象，外链建设人员辛苦交换友情链接几个月，还不如高层说的几句话效果好。

17．请求链接

如果网站上真的有非常有价值的内容，直接请求链接也是一个可行的方法。虽然大部分商业性质网站不会轻易链接到其他网站，但还是有很多科研机构、学校、政府部门网站及个人博客，会在页面上列出对用户有帮助的资源。尤其是如果对方网站上已经有现成的资源列表，写一封诚恳的邮件，告诉对方自己网站上有很好的内容，能增强对方内容的完整性，不少站长会同意加上你的链接。

18．购买网站

如果碰到不错的外链来源网站，但对方不同意链接，也可以考虑把对方网站购买下来。这在某些竞争非常厉害的行业是个常见的手法，比如旅游、成人、药品、法律服务等类网站。

寻找购买对象不一定必须要求排名非常好。排名在前十位、二十位的网站不会轻易出售，要求的价格也一定比较高。可以找排名在前十页或二十页的网站，其实都有不错的权重和相关性，不然也不会战胜其他几十万几百万竞争对手出现在前十页或二十页。不过出现在前十页，关键词点击流量微乎其微，对原来的所有人来说很可能是个鸡肋，报上一个不错的价格，也许对方就能卖给你。

19．站群

与购买网站有点类似的是通过站群建设外部链接，这又是一个有争议的灰色地带。拥有多个网站本身不一定是作弊或黑帽，但是超过一定程度被判断为作弊的可能性将大大提高。

通过站群建设外部链接要注意下面几点。

（1）数量不要过多。拥有十几个二十个网站很正常，但一个小公司或个人站长拥有几百个同行业网站就不正常了。

（2）站群之间不要交叉连接。使用站群的目的是集站群的力量，共同推一个主网站。所以从站群网站链接到主网站更为恰当，这些站群之间不应该交叉链接。

（3）不要使用重复内容。站群里的每个网站都应该有自己的独特内容，使用一样内容建多个网站毫无意义，对用户没有帮助，搜索引擎也不会喜欢。

（4）网站主题不同。虽然应该尽量保持站群与主站的内容相关性，内容应该是相同或相近行业，但不同网站之间还应该有一定的主题差别。比如主网站是减肥药物，站群可以分别以健身、美容、减肥方法、养生等为主题。

（5）可能的情况下站群网站放在不同的服务器和 IP 地址上。

（6）可以被接受的站群与黑帽站群之间的差别就在于用户体验。建立站群时应该问

问自己，一个新的网站是否能够为用户至少提供一定的独特价值？还是纯粹为了多些外部链接？

20．赞助活动

赞助活动也是一个建立单向外部链接的不错方法。公司赞助各种形式的会议、公益活动，通常活动官方网站上会列出赞助商并且给予链接。有的政府或行业机构网站很难获得链接，赞助他们举办的活动几乎是唯一的门票。

当然赞助活动一定有成本，不过有的时候成本不一定很高。除了现金赞助，也可以是产品赞助。某些产品成本并不高，像软件、咨询服务等。甚至赞助活动可能以优惠券、打折券就能实现，既能提高曝光度、带来外部链接，还能带来直接销售，非常划算。

21．电子书发布

如果你有比较全面深入的针对某个主题的内容，可以做成电子书，书里自然可以放上作者网站链接。正文中需要引用和进一步解释时，还可以链接到网站的内页。

电子书最常见的格式是 PDF，也可以是普通的 Word 文件，甚至是 PPT 格式，文件中都可以留下链接。除了发布者自己的网站上提供下载，电子书还可以自行提交到各大软件下载网站。很多用户喜欢传播电子书，只要在自己的网站和几个主要软件下载网站上发布，就会迅速在网上传播开来。

我在自己的博客上发布的几本电子书，都没有做任何其他形式的推广，就有很多站长把电子书放在自己网站上提供下载。电子书中的链接也会被搜索引擎当做外部链接，如图 6-42 所示。

图 6-42　PDF 文件中的链接

22．信息发布网站

网上有一些允许用户自由发布信息的网站，如 B2B、分类广告等。其中一部分允许用户在发布信息中包含链接。

本节简单讨论了 20 多个常规外链建设方法。这些方法全部使用既不现实，也没必要。外链建设人员尝试各种方法后，最好选择几种自己最擅长的，持之以恒，积累经验和资源。简单方法做到极致，就是最有效的方法。

6.11 非链接形式的链接

前面讨论的都是正常意义上的、写在<a href=>标签里、连向网页的链接。

网上也有很多非正常链接形式的 URL 出现，这些在正常意义上来说隐藏着的链接，是否会对搜索引擎发现页面、传递权重、页面排名有影响呢？目前谁也说不准。

比如 Gmail 中出现的链接都是被 Google 跟踪的。点击 Gmail 里的链接，你会在地址栏中首先看到跟踪代码，然后再转向到真正的 E-mail 中出现的 URL。到目前为止，没有迹象表明 Gmail 里的链接会对页面收录有什么帮助，更别说对排名的影响了。不过，如果 Google 从来没想过利用这个数据，为什么从一开始要跟踪呢？

非一般的、隐藏形式的链接（准确地说只是可能的链接）还包括：

（1）其他网站引用你网站上的图片文件，也就是图片盗链。网站 B 使用网站 A 的图片，HTML 代码为<img src="http://www.domainA.com/images/pic.jpg">，其中的 http://www.domainA.com/images/pic.jpg 算不算对域名 A 的一个链接？

（2）链接到 JavaScript 文件。<a href="http://www.domainA.com/script.js">算不算对域名 A 的一个链接？姑且不去讨论为什么要链接到一个 JS 文件。

（3）链接到 CSS 文件。<a href="http://www.domainA.com/style.css">算不算对域名 A 的一个链接？

（4）连到 RSS 种子的链接。<a href="http://www.domainA.com/rss.xml">算不算对域名 A 的一个链接？

（5）搜索引擎能看到的 E-mail 里的链接，诸如雅虎邮箱，Hotmail，Gmail。

（6）在 JS 脚本或 JS 注释里出现的 URL。

（7）在 CSS 或 CSS 注释中出现的 URL。

（8）图片、视频文件 meta 数据中出现的 URL。

（9）HTML 文件注释里出现的 URL。

（10）HTML 文件头部，meta 数据，以及 ALT 标签等地方出现的 URL。

（11）可以被工具条跟踪、用户访问的，但没有出现在其他页面的 URL。

（12）需执行 JS 脚本后才能看到的目标 URL。搜索引擎都在尝试解析 JS 中的链接，从而发现新页面，其中的 URL 算不算链接？能否传递权重？

（13）需登录才能看到的链接。有的付费内容网站可以通过所谓“First Click Free”机制，允许 Google 抓取需登录才能看到的内容和链接。

（14）Flash 中的链接。

（15）纯文字、没放进<a>中的 URL，如页面上仅仅出现文字 http://www.chinamyhosting.com/，不是可以点击的链接。

（16）非网页文件，如 TXT 文件中出现的链接。

（17）加了 meta noindex 标签的页面上的链接。

（18）域名注册数据和 DNS 数据。

（19）可提交的表单里出现的链接。

（20）小工具软件里的链接。

（21）广告链接如百度推广、Google Adwords 指向的链接。

最后提一下图片中可能出现的链接。我们都知道搜索引擎是看不懂图片的，这也是做 SEO 的都强调文字内容的原因所在。但 Google 2006 年就收购了一家图像识别处理公司 Neven Vision，最近又在鉴别图片颜色、类型、面部相似度等领域不断有新花样。那么能否从图片中识别 URL 呢？如果可以，Google 地图街景照片中出现的广告 URL 能否被读取呢？这可能也是 Google 在研究的方向之一。

上面提到的这些 URL 是否有助于搜索引擎收录新页面？是否会在某种程度上传递权重？对排名有什么影响？SEO 界没有定论。以后如果有进一步信息，我会在 SEO 每天一贴博客中分享。

6.12 竞争对手能否通过垃圾外链陷害你

网上搜索排名竞争日趋激烈，有些不道德的站长无法把自己的网站优化好，开始想歪主意陷害竞争对手。

他们要想陷害别人，无非从两方面入手：网站内和网站外。DDOS 攻击，黑入对方网站加上隐藏文字、隐藏链接、非法内容，甚至修改 robots 文件禁止收录，显然可以陷害竞争对手。不过这是与 SEO 无关的违法行为，在这里不去讨论。

竞争对手能否通过外部链接陷害你呢？前面讨论过，外部链接要做到内容相关、自然平稳增加。如果竞争对手刻意给你的网站制造一些不相关、快速增加、不自然的垃圾链接或购买的链接，能否使你的网站排名受到影响？

早在 2005 年就有人注意到，在某些极特殊情况下这种陷害方法是可以成功的，称为“Google 保龄”（Google Bowling），取其“用保龄球把其他人排名打倒”之意。

可能的 Google 保龄方法包括：

- 给竞争对手购买大量全站链接，比如买卖链接最常见的页脚全站链接，使对手网站一夜之间多出成千上万的外部链接。
- 从典型的链接工厂给竞争对手制造链接，锚文字全部一样。商业性热门关键词作为锚文字，可能产生的效果最大。
- 给竞争对手制造大量桥页，页面充满关键词，机器生成内容，还可以用上各种作弊方法，如隐藏链接、隐藏文字、关键词堆积，然后链接向竞争对手。
- 制造大量垃圾页面或违法内容，然后用 301 转向、JS 转向、Meta 转向等转向到竞争对手网站。
- 制造大量垃圾链接，如群发博客评论、留言本等。

按理说，竞争对手在网站之外创造垃圾内容和垃圾链接，不应该影响你的排名，站长无法控制网站之外的内容，也无法阻止竞争对手陷害。在 Google 站长帮助文档中有一句很重要的话，中文版内容如图 6-43 所示。

对应的英文版如图 6-44 所示。

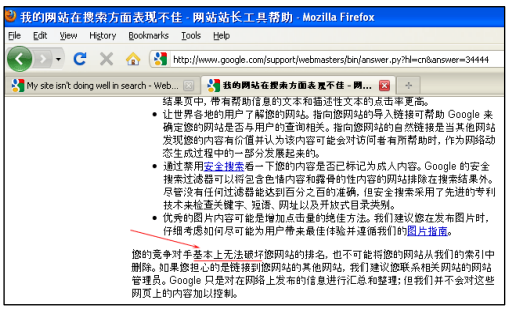


图 6-43　Google 站长帮助文档

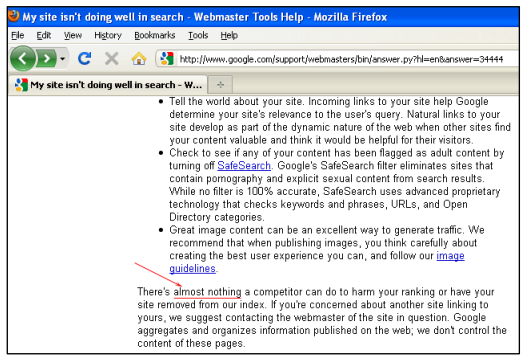


图 6-44　Google 站长帮助文档英文版

可以看到中文版翻译并不太准确，准确的翻译应该是：

您的竞争对手基本上无法破坏您网站的排名，或者将您的网站从我们的索引中删除。

抓图中箭头指向关键的用词——“基本上无法破坏”，换句话说：在极特殊的情况下还是可以破坏的。有不少人观察过 Google 保龄的现实案例，我本人前不久也看到过被竞争对手制造大量黑链，导致网站被惩罚很长一段时间的例子。所以无论从原理上还是现实中，竞争对手通过外部链接破坏你的排名是可能出现的，虽然极罕见。

虽然称为“Google 保龄”，但这种现象不限于 Google，同样也出现在百度等搜索引擎。

要防止这种陷害，唯一的方法是提高自己网站的权重。Google 保龄案例都是发生在新站小站上，还没有看到过真正的权威网站被竞争对手陷害成功的案例。Google 保龄只有在网站自身没有比较强的外部链接情况下才会出现，所以只要你花时间做好、做强自己的网站，外部链接构成健康，有大量高质量外部链接，锚文字自然分散，竞争对手就没办法陷害。

一般来说，SEO 不必为竞争对手通过外部链接陷害担心。当你的网站没权重、没排名时，没人会陷害你；当你的网站有权重、有排名时，陷害是很难的。

6.13 链接工作表

一般公司网站都应该有页面修改更新日志，方便日后查看哪些改变引起了什么效果。除此以外，外部链接建设也应该有一个工作表，方便统计记录工作进度、多个网站之间共享数据及新员工的工作交接等。

链接工作表至少应该包含以下内容（如表 6-2 所示）：

- 对方网站名称。
- 首页 URL。
- 首页 PR 值。
- 主题匹配度。SEO 人员给目标网站大致做个判断，给予高、中、低评级。相关度越高，需要投入的时间精力分配越多。
- 联系方式。可以是对方站长 E-mail 地址、提交表格地址、QQ、MSN 账号等。
- 链接页 URL。对方给予链接的具体页面，不一定是首页。
- 链接页 PR。首页 PR 高，并不意味着链接页 PR 值也高。有的时候由于网站结构的问题，给出链接的页面甚至可能没有被收录，这样的链接就没有效果。
- 链接种类。可以是交换链接、博客文章发布、论坛签名、目录提交等。
- 链接使用的锚文字。
- 本站友情页面。如果是交换链接，列出对方网站在自己网站上的链接 URL。
- 第一次请求日期。注册提交或第一次联系友情链接的日期。
- 跟进日期。对方没有回复一段时间后，跟进联系日期。
- 状态。可以是已完成，已放弃，处理中等。
- 备注。其他相关信息，如付费目录费用。
- 处理人。

表 6-2　链接工作表

网站名称	首页URL	首页PR	主题匹配	联系方式	链接页	链接页 PR	种类	锚文字	本站友情页面	请求日期	跟进日期	状态	备注	处理人

最简单的链接工作表可以是一个 Excel 文件。有技术实力的公司也可以把链接工作表做成软件或线上工具形式，方便团队沟通、协作。除了上面列出的内容外，还可以增加一些功能，如：

- 友情链接认证。软件自动访问对方链接页，看自己的链接是否还在。
- 计算工作量。根据处理人、联系日期、跟进日期及链接状态，自动计算外链建设人员的工作进度和工作量。
- 对方网站质量。自动查询线上工具，列出对方网站的收录数、域名注册日期等。
- 惩罚可能性。搜索对方网站名称、主要目标关键词，记录排名情况。如果搜索网站名称网站不在第一位，就可能是被惩罚的网站。

备选网站。根据第 2 章中讨论的特殊搜索指令，在搜索引擎中找到可能的链接来源，供外链建设人员人工联系和跟进。

第 7 章

SEO 效果监测及策略修改

效果监测是 SEO 项目的重要步骤之一。SEO 是一项不能停的工作，效果监测既是前一轮 SEO 的总结，也是下一轮 SEO 的开始。

7.1 为什么要监测

再漂亮的计划，再有力的执行，都有改进的空间，而不被测量的东西是无法着手改进的。

7.1.1 检验工作成效

SEO 监测的意义当然首先在于检验 SEO 成效，不仅向公司高层、其他部门汇报 SEO 对网站流量及赢利的贡献，也使 SEO 团队本身了解自己是否走在正确的方向上？所做的工作完成了多少？到底对网站有多大贡献？

要比较准确地统计工作成效，必须设定监测基准。在 SEO 实施之前就需要设定好哪些指标需要监测，网站实施 SEO 前各指标的表现数据。虽然效果监测是在第一轮 SEO 工作完成之后才能具体进行和分析，但监测事项必须在 SEO 开始实施前就做好规划。没有记录网站原始表现，到后面就无法判断 SEO 效果。

另外，为了更准确地知道 SEO 成效，还要同时监测主要竞争对手情况，如下面提到的应该监测的排名、流量、收录、外部链接等，只要是可以查询到的数字，都要进行周期性记录。SEO 不是自己网站单独能够体现出效果的，竞争对手也同时在优化、提高，只看自身数据，很可能无法完整了解 SEO 是否达到了应有的水平。

7.1.2 发现问题，修改策略

SEO 计划和实施是否真的符合搜索引擎质量规范和算法规律，谁也不能事先百分之百确定。很多细节问题 SEO 行业内部都没有共识，到底每一步、每一个元素是否优化得当，自己说了不算，只能以最后效果为依据。实施 SEO 一段时间后，检查各项指标、分析流量，才能发现 SEO 过程中可能存在的问题并进行策略修正。

在这个意义上，SEO 是个不间断的调整过程。当初的计划可能有偏差，很多细节一定有不完善的地方，竞争对手也在加强网站，搜索引擎算法同时在不断调整中，因此，SEO 策略修正是必不可少的步骤。

7.1.3 SEO 完整过程

从前面几章的讨论可以得出 SEO 的完整过程。

- 竞争研究：包括关键词研究和竞争对手研究。
- SEO 计划：诊断网站，找出不足之处，提出优化方案及外部链接建设方案。
- 网站优化：包括结构调整及页面优化。
- 外部链接建设：通常与网站优化同时进行。
- 效果监测及流量分析：本章讨论的内容。
- 策略修改：基于监测数据，必要时重复从关键词研究开始的整个过程。

所以，SEO 是一个循环往复的过程，效果监测和策略修改是其中重要的一步。

7.2 网站目标设定及测量

SEO 是网站营销活动的一部分，必须帮助达成网站目标。而不同的网站目标对 SEO 策略也有影响。

7.2.1 网站目标

做网站时，很多企业和个人站长经常会忽视的一个步骤是明确定义网站目标。每个网站的存在都有它的原因，但很多站长说不清网站的目标。有的站长认为建站就是为了赚钱，当然这可以是建站的最终目的，却不能成为营销意义上的目标。我们需要再进一步明确：网站怎样赚钱？

网站目标必须是实在的、可以操作的、用来指导网站设计及所有网站运营行为的，而不是笼统的、表面的、敷衍了事的。

网站目标必须在动手建网站之前就确定下来，而不能走一步看一步。网站建设、设计过程中遇到多种选择时，取舍的标准就是看是否有利于达成网站目标。

这里所说的网站目标是给企业和站长本身看的，而不是给用户、投资人或老板看的。在一些网站项目策划书或网站页面上，我们经常看到网站目标是诸如“给客户提供最大价值”或“为客户提供完整解决方案”。这是给客户看的，同时也是不着边际、无法操作的。有的可能说网站目标是“促进公司在线销售增长 20%”，这是给老板看的，对运营网站也没有直接指导意义。

网站目标要非常明确，具有可操作性。但是很多时候没有经验的人不一定能轻易正确定义网站目标。最常见的错误就是把网站目标定得过于宽泛，无法指导实践。比如认为网站目标就是赚钱，或者就是促进销售，或者是建立网络品牌，这些都是看似正确、却不能指导网站实践的很宽泛的目标。

对网络营销人员来说，网站目标就是你想让浏览用户在网站上做什么。你最想让浏览者做的那件具体的事就是你的网站目标。

7.2.2 网站目标实例

网站目标有时候并不像说起来那么简单明显。大部分电子商务网站最容易想到的网站目标是让用户把产品放入购物车，最终实现购买。但实际上每一个网站都应该以直接产生销售为目标吗？不一定，不同的网站应该找出最适合网站产品、用户特征、业务流程的目标。下面举几个具体的例子，帮助站长明确怎样定义网站目标。

用户初次访问一个陌生网站很少有立即购买产品的。据研究，用户与网站接触 5～7 次才达到较高的购买率。但除非网站是知名品牌，一般浏览者一旦离开网站很少会回来，怎样与潜在用户保持接触？电子邮件营销是最好的方法。如果网站运用了电子邮件营销，那么网站目标很可能应该是尽最大可能让浏览用户注册电子杂志，收集尽可能多的潜在客户的电子邮件，而不是试图说服初次访问者买东西。

再比如，如果销售的产品是软件，很可能网站目标不是让浏览者直接购买，而是吸引浏览者下载免费试用软件。最后的销售达成是通过软件本身的易用性及强大功能，试用软件有效期结束后用户需要付费才能继续使用，或者限制免费版功能。在网站上，站长最想让浏览者采取的行动就是下载试用软件，这就是网站目标。免费下载也可以是电子书、行业报告、音频视频文件、屏幕保护、图片等。

有的网站是为了最终提供某种无法线上完成或必须客制化的服务，这时网站目标就可能是促使浏览者与企业联系，了解进一步情况。联系的方式可能是填写在线表格，可能是发电子邮件，也可能是直接给企业打电话。无论是哪种联系方式，站长都应该很明确，网站目标不是直接产生线上订单，而是促使浏览者采取行动，联系企业或站长。

有的网站延续线下营销模式，并不试图直接产生销售，而是希望用户索取样品并使用样品，然后再通过后续联系，或者业务人员直接推销等方法产生最终销售。这时网站目标就应该是促使浏览者填写表格，要求网站寄样品或产品目录。

有的网站是为了辅助实体店的营销活动，网站本身并不能产生任何交易。当然对有实体店的企业来说，线下与网上直接销售并不矛盾。但有时企业由于某种原因建设网站纯粹是为了辅助实体商店，这时网站目标该怎样定义需要花费一点心思。比如，网站目标可能是促使用户拿起电话，询问实体商店的具体地址、营业时间。如果再进一步考虑，是否还能有更具吸引力的网站目标？比如在网站上提供专用折扣券，浏览者在网上打印折扣券，可以到实体店使用。这时网站目标就可以很明确地定义为吸引用户打印折扣券。

有的网站主要依靠广告赢利，网站目标显然就会与上面讨论的以销售产品或服务为主的网站不同。如果网站是以 PPC 广告，也就是按点击付费广告为主，那么网站目标就

是让浏览者点击广告链接。网站给出的其他选择越多，浏览者点击广告的可能性就越低。所以在设计网站时就要考虑怎样在不产生欺诈的情况下，吸引用户点击广告，甚至有时候可以让用户别无选择，只能点击广告离开网站。

如果网站是以 CPM，也就是按显示次数付费为主的广告，策略则刚好相反。你需要浏览者在网站上停留时间越长越好，打开页面越多越好，最好不要离开网站。

不同的网站赢利模式、产品特性、用户行为方式、业务流程，会形成不同的网站目标，站长需要潜心研究才能正确地定义网站目标。

7.2.3 网站目标确定原则

好的网站目标应该具备下面几个特点。

第一，**现实性**，也就是说用户有可能达成。不要给用户增加难度，设定一些需要克服困难才能达成的网站目标。如果网站销售的是价值几十万、几百万的工业设备，网站目标就不要设定为直接在线销售，而是设定为让浏览者联系企业，或者询问进一步情况，或者索取报价单、产品目录，或者直接打电话给销售部门。总之，不要让用户在网上直接订购这么大金额的产品，这往往是不现实的。

第二，**可测量性**。网络营销相比传统营销最大的优势之一就是营销效果可以精确测量和分析，在此基础上改进营销手法。网站目标的设定也同样如此。站长想让浏览者做的事，应该是可以通过某种方式测量的。有时候测量方法简单明确，比如产生在线订单，或注册电子杂志，下载软件、电子书等，这些都会在网站后台或服务器日志文件中留下明确数据。

有的时候则要想一些方法才能使网站目标的测量可以实现。比如，上面所说的吸引用户打印折扣券的情况。用户打开折扣券所在页面次数可以测量，但有没有打印，站长是不容易知道的。要想测量这个目标，就要结合几种方法，比如折扣券所在网页的浏览次数，折扣券在实体店被使用的次数，如果技术条件允许，可以动态生成折扣券系列号码等。

第三，**行动性**。网站目标应该是特定的用户行为，也就是用户要做的一件具体的事情。像上面所列出的例子，无论是生成订单还是注册电子邮件列表，或者下载，填写在线联系表格，打电话，点击链接等，这些都是明确的用户行为。有一些似是而非的指标就不是真正的网站目标，如网站流量增长 10%，或转化率提高 1%等。这些可以是用在企业展望上的目标，却不是设计、运营网站时用于指导实践的网站目标。

第四，**明确单一**。如前面提到的例子，网站目标都是很明确单纯的。网站目标设定不能多，不然在建设网站时遇到需要做选择决策的问题时，可能会产生冲突，难以抉择，不知道让浏览者做这个好，还是做那个好，网站设计也就容易重点不突出。

单一目标是最有力的。在某些情况下单一目标达成困难，可以加一个辅助目标，如购物类网站，第一位的网站目标当然是直接产生销售，如果浏览者没在网站买任何东西，就想办法让他们注册电子杂志，作为辅助网站目标。除非是很大型、功能复杂的网站，不然千万不要面面俱到，既想让浏览者做这个，又想让浏览者做那个，结果可能是哪个都没达成。

7.2.4 网站目标影响 SEO 策略

SEO 人员为什么必须清楚了解网站目标？最重要的是，网站目标影响 SEO 策略。

例如，信息类网站靠显示广告赢利，网站目标是尽量增加页面访问数（PV），电子商务网站，以在线销售为目标，SEO 策略将产生很大不同。电子商务网站希望的是能转化的高质量目标流量，关键词研究的重点在于挑选和优化最相关的、交易类的关键词。信息网站则对流量相关性要求不高，只要是流量就能增加 PV，关键词研究重点在于发现新热点、不断扩展长尾。

在网站结构上，要提高 PV，将文章分页是方法之一，所以大家可以看到很多门户、资讯网站将明明不长的文章也分为几页。代价是，页面和链接权重分散，收录可能成为问题，甚至可能影响用户体验。

另外，网站目标的正确设定和测量直接影响 SEO 效果的监测及策略修改。SEO 效果不仅表现在排名和流量上，更表现在带来转化上。

7.3 非流量数据监测

SEO 需要监控的数据可以分为非流量和流量两部分。本节讨论非流量数据，下一节介绍流量分析。

7.3.1 收录数据

收录是网站排名和流量的基础，尤其是大中型网站。优化不好的网站，经常收录不充分。SEO 团队要解决的一个重大问题就是，使尽量多的页面被搜索引擎收录和索引。下面几个收录数据需要记录跟踪。

1．总收录数

基本做法是使用 site：指令查询搜索引擎对某个网站的总收录页面数，再加上站长自身知道的网站实际页面数，计算出收录比例。优化好的网站收录率能达到百分之七八十。

不过 site：指令往往并不准确，尤其是 Google，近一两年 site：指令越来越不准确。为了得到比较准确的收录数，可以尝试使用 site:domain.com，加上网站每个页面都会出

现的词或句子，比如网站名称、出现在页面顶部的口号、电话号码、地址、备案号等，如图 7-1 所示。

图 7-1　site:域名+网站名称

需要注意的是，出现在 HTML 底部的文字有可能并没有被搜索引擎索引。有的页面文件太大，搜索引擎虽然抓取整个页面 HTML 代码，但不一定把页面上所有文字进行索引，而只是索引前半部分。所以如果页面文件比较大，需要搜索页面上靠前的独特文字，才能得出较准确的收录数字，通常位于底部的备案号可能不适合。

Google 收录数也可以参考 Google 网管工具中指向首页的内部链接数。通常网站所有页面都有首页链接，所以首页的所有内部链接数也就是收录的页面数。当然，这个方法只限于 Google，也仅限于查询自己网站。

选定一个查询方法，就坚持使用这个方法和指令。在不同指令的收录数字间比较是没有意义的。

2．特征页面收录

除了网站首页外，再从分类页面中选一部分有特征的或典型的页面，以及产品或文章页面中的一部分典型页面，查看这些典型页面是否被收录。大中型网站可能需要记录上万个特征页面。在选择典型产品页面时，既要兼顾到尽量多的分类，也要兼顾到不同时间发布的产品，既要有最早发布的、因此已被推到离首页比较远的页面，也要有比较新的、离首页比较近的页面。

查询特征页面收录与否，只要在搜索引擎输入这个页面的完整 URL，没有结果就是没有收录。

检查特征页面收录情况，经常能快速、直观地判断出网站哪些部分获得的链接少、权重太低。

3．各分类收录数

每个分类下的产品页面收录数是多少？记录这个数据有助于了解哪些分类收录完整，哪些分类由于内部链接结构的问题收录不充分，并采取对应措施。

各分类下的页面标题标签和 URL 格式比较规则时，SEO 人员可以灵活运用 site:指令及 intitle:、inurl:等组合，得到各分类的收录数。例如如果页面标题是第 5.1.5 一节推荐的标准格式：

产品名称/文章标题–子分类名称–分类名称–网站名称

使用指令：

site:www.domain.com intitle:分类名称

就可以得到这个分类的页面收录数。

如果页面 URL 比较规则，分类首页 URL 为：

http://www.domain.com/catA/sub-cat-1/

产品页面 URL 为：

http://www.domain.com/catA/sub-cat-1/page-1.html

产品页面都处于上级分类目录下，使用指令：

site:www.domain.com/catA/sub-cat-1/

或：

site:www.domain.com inurl:/catA/sub-cat-1/

就可以得到子分类 1 的页面收录数。

如我的博客，2009 年 5 月有 20 篇帖子，如图 7-2 所示。

Google 收录 21 页，如图 7-3 所示。

图 7-2　2009 年 5 月实际帖子数为 20

图 7-3　Google 收录 2009 年 5 月页数为 21

百度收录是 25 页，如图 7-4 所示。

图 7-4　百度收录 2009 年 5 月页数为 25

收录页面包括按时间分类等，所以超过实际帖子数。大致可以判断，2009 年 5 月份的帖子收录不错。

如果各分类下页面标题、URL 等都不规则（某些情况下这可能说明网站结构规划有问题，本应该尽量避免），可以采取抽样统计方法，选取每个分类下一定比例的页面，检查收录与否，进而计算出这个分类下页面的收录比例。这可以与统计特征页面结合起来做。

从前面几章的讨论可以看出，网站收录不充分，经常是下面几个原因：

- 域名权重不够高，很多内页权重降到收录最低要求线之下。
- 网站结构有问题，使搜索引擎无法顺利爬行。
- 内部链接分布不均匀，使某些分类权重不够高，这个分类下的大量页面不能被收录。
- 搜索引擎不友好的因素，如 JS 链接、框架结构、大量使用 Flash 等。
- 网站内容原创度不够，大量转载和抄袭内容搜使索引擎认为没有必要收录。

跟踪、记录网站页面收录情况，使 SEO 人员能够知道网站结构调整后，是否达到提高收录率的目的，如果没有，应该仔细检查上述几方面。

收录数据的局限在于，收录数高并不一定意味着流量高（虽然大部分情况下是如此）。有的页面即使被收录，由于权重太低或页面优化不好，也没有任何排名，带不来流量。为了矫正这个偏差，SEO 人员也可以记录网站有效收录页面数，也就是带来过至少一个搜索流量的页面数。

GA“最常见目标网页”部分，在右上角选择流量细分中的搜索流量，如图 7-5 所示。

GA 将显示所有带来过搜索流量的页面总数，如图 7-6 所示。

其他页面即使已被收录，却没能带来哪怕一个搜索流量，可以在一定程度上理解为无效收录。GA 显示的数字来源于真实流量，不依赖于 site:之类指令的算法可靠性，所以数字是比较准确的。

图 7-5　选择来自搜索流量的最常见目标网页

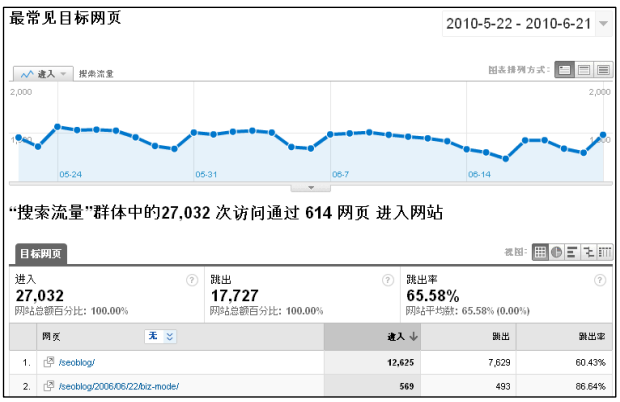


图 7-6　有效收录页数

7.3.2　排名监测

所有 SEO 都不会忘记的肯定是监测关键词排名。一般来说，为了完整地了解网站关键词排名情况，需要监测下面几种关键词：

- 首页目标关键词。
- 典型分类页面目标关键词。
- 典型最终产品或文章页面关键词。

视网站大小，典型分类和最终内容页面可能需要监测几十个甚至几百个。最好每天记录排名，或至少每周固定时间记录排名。需要记录的关键词比较多时，人工查询不太现实，肯定需要关键词排名查询工具。后面 SEO 工具章节介绍的追词、观其关键词排名工具都不错。

监测关键词排名的局限在于，排名与流量不一定是直接对应关系。很多时候排名在

前面，并不意味着能带来搜索流量，尤其是没有进行关键词研究，或者关键词研究方法错误时。

有一些 SEO 服务商还刻意误导客户，把关键词排名作为 SEO 的唯一指标，挑选目标关键词时，又选择那些没有人搜索，当然也没有难度的词。这样的关键词排到前面毫无意义。

另外，近几年搜索引擎大量引入个性化排名和地域性排名，使得 SEO 人员自己查到的排名与众多用户实际看到的排名不一定相同。这里并不存在作弊或误导，只不过是因为搜索引擎为增强用户体验引入的功能。

所以排名只能作为 SEO 指标的一部分，绝不是 SEO 的最终目的。现在国外一些 SEO 服务商已经不再提供排名报告给客户。

7.3.3 外部链接数据

外部链接增长情况也是衡量 SEO 工作成效的重要部分。需要记录的数据包括：

- 首页总链接数。
- 网站所有页面的总链接数。
- 连向网站的总域名数。
- 特征页面链接数。

搜索引擎和几乎所有外链工具并不实时更新外部链接数字，所以每天查询没有什么意义，能做到每周甚至每个月查询一次就足够了。记录下不同时间的外部链接数，可以大致看到：

- 随着网站内容建设和外部链接建设的进行，外部链接总体增长趋势如何？
- 哪个链接诱饵计划、实施得当？
- 哪个新闻或公关活动成功？
- 与排名或流量升降是否有时间和因果上的关系？
- 竞争对手外链有没有突然增减？为什么？有什么可以借鉴的？

查询外部链接使用搜索引擎自身的 link:指令非常不准确，百度甚至根本不支持这个指令。目前比较可靠的外部链接查询有三个方法：Yahoo! Site Explorer，SEOMoz 的 Open Site Explorer，Majestic SEO，在 SEO 工具一章都有详尽介绍。

Google 网管工具也列出网站的外部链接，还提供数据下载。缺点是，其数据不定时更新，没有什么规律，而且只能查看自己的网站。

7.3.4 转化和销售

无论是收录、排名和流量，最终都是为了转化。记录网站每个时期的转化和销售数

字，可以更直观地看到 SEO 给网站带来的实际好处。

除了记录转化次数，还可以根据平均每单销售额计算出每一个独立 IP 的价值。如网站转化率为 1%，100 个独立 IP 带来一个销售，平均每单销售额为 50 元，那么一个独立 IP 的价值就是 0.5 元。给流量赋予一个金额数字后，SEO 成效就能体现在非常具体的、给公司带来的销售额甚至利润上。

从后面流量分析一节还可以看到，转化及销售额、利润都可以和具体流量来路（搜索流量、直接输入流量、来自其他网站的点击流量）以及搜索流量中是哪个关键词对应起来。这样，什么样的流量价值最高，什么关键词价值最高，通过网站流量分析一目了然。

在 SEO 项目中监测转化与销售数字的局限是，转化与销售并不受 SEO 控制，更多地是受制于产品本身质量、价格、文案写作、客服水平、品牌知名度等。

7.4 流量数据监测

大部分情况下，流量分析要借助于流量分析软件，但直接读网站日志还是站长及 SEO 的基本功之一。

7.4.1 怎样读日志文件

网站服务器会把每一个访问信息、每一个服务器动作、每一个文件调用自动记录下来，存在服务器原始日志文件中。所以，日志中的信息是相对准确、全面的。有些 SEO 需要知道的信息在流量分析软件中很少支持，必须直接查看日志，如服务器返回的状态码。

原始日志就是一个纯文本文件，只要用文字编辑软件如 WordPad 或 Notepad 打开就可以。一般主机商会在控制面板提供日志文件下载。

下面是从我的博客“SEO 每天一贴”（原网址为 http://www.chinamyhosting.com/seoblog/，2011 年初换到另一个域名 http://www.seozac.com）2010 年 8 月日志文件中随机选取的一行，我们来看一下它包括哪些信息：

```
121.11.255.42 - - [31/Aug/2010:20:44:36 +0800] "GET /seoblog/ HTTP/1.1"
200 52778 "http://www.baidu.com/s?wd=seo" "Mozilla/4.0 (compatible; MSIE
6.0; Windows NT 5.1; SV1; 360SE)"
```

用户 IP 地址：

```
121.11.255.42
```

这是访问用户所在的 IP 地址，显示出访问的人来自什么地理位置。在 IP 地址信息

服务查一下这个 IP 所属位置，可以看到这位访客来自广东（如图 7-7 所示）：

http://www.123cha.com/ip/?q=121.11.255.42

图 7-7　IP 地址显示用户地理位置

日期/时间：

```
31/Aug/2010:20:44:36
```

这是文件被访问的准确时间。和 IP 地址结合起来，查看多条日志记录就可以跟踪某一个特定的用户从一个网页到另一个网页的访问顺序和在网站上的活动。

时区：

```
+0800
```

相对格林威治时间的时区差。中国、新加坡处于东 8 区。

服务器动作：

```
"GET /seoblog/ HTTP/1.1"
```

服务器要做的动作要么是 GET，要么是 POST。除了一些 CGI 脚本外，通常都应该是 GET，也就是从服务器上获取某个文件，可以是 HTML、图片、CSS 等。

例子里的这段记录意思就是，按 HTTP/1.1 协议获取 URL /seoblog/ 处的文件。这里的 URL 是相对地址，已经省去了域名部分，所以/seoblog/ 指的就是博客首页 http://www.chinamyhosting.com/seoblog/。

上面的例子是访问一个页面时的记录。在日志中，每一个图片、JS 脚本等文件的访问也都会有一行记录。如：

```
58.63.31.46 - - [31/Aug/2010:21:52:23 +0800] "GET
/seoblog/wp-content/themes/mistylook/style.css HTTP/1.1" 200 9450
"http://www.chinamyhosting.com/seoblog/" "Mozilla/4.0 (compatible; MSIE
6.0; Windows NT 5.1; SV1; CIBA; .NET CLR 2.0.50727; 360SE)"
```

这行日志代表获取的是博客模板的 CSS 文件。

服务器状态码：

```
200
```

服务器返回的状态码。200 指成功获取了文件，一切正常。如果返回 404，就是文件不存在/没有找到。其他常见状态码包括：

- 301——永久转向。
- 302——暂时转向。
- 304——文件未改变，客户端缓冲版本还可以继续使用。
- 400——非法请求。
- 401——访问被拒绝，需要用户名、密码。
- 403——禁止访问。
- 500——服务器内部错误，通常是程序有问题。
- 503——服务器没有应答，如负载过大等。

文件大小：

```
52778
```

指的是所获取文件的大小，例子中是 52 778 字节。

来路：

"http://www.baidu.com/s?wd=seo"

显示访问者是从哪里来到当前网页，也就是来到这个网页之前访问的那个网页 URL。来路可能是同一个网站的其他页（用户通过点击网站内部链接浏览），有可能是其他网站（用户通过其他网站上的链接点击过来），也有可能是搜索引擎的结果页面，如上面所示的例子。

来路是搜索引擎结果页面时，URL 中含有一个对 SEO 意义重大的信息，就是用户搜索的关键词。例子里的来路是百度搜索，搜索关键词是“seo”。

浏览器类型（User Agent）：

```
"Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; 360SE)"
```

最后一段显示的是浏览器和用户计算机的一些信息。

例子中这段信息表示用户使用的是：

- 与 Netscape 兼容的 Mozilla 浏览器。实际上大部分浏览器 User Agent 都使用 Mozilla/为开头，所以并不能区分浏览器。
- 浏览器是 IE6.0，或 IE6.0 为内核。

- Windows NT 操作系统。
- 360 浏览器（360 浏览器是以 IE6 为内核的）。

如果用户使用的是其他类型计算机或浏览器，在这一段还可能看到这类代码：

- Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0; Maxthon; Alexa Toolbar)——傲游浏览器。
- Mozilla/5.0 (Windows; U; Windows NT 5.1; zh-CN; rv:1.9.2.8) Gecko/20100722 Firefox/3.6.8 GTB7.1——火狐浏览器。
- Mozilla/5.0 (Windows; U; Windows NT 6.1; en-US) AppleWebKit/533.4 (KHTML, like Gecko) Chrome/5.0.375.127 Safari/533.4——Google 的 Chrome 浏览器。
- Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1;) ——微软 IE6 浏览器。

搜索引擎蜘蛛就相当于一个浏览器，第 2.4.1 节中提到的搜索引擎蜘蛛用于表明身份的用户代理就是这段信息。如百度蜘蛛如下表明自己：

```
Baiduspider+(+http://www.baidu.com/search/spider.htm)
```

日志是网站访问的最真实记录。分析用户访问时还可以借助下面介绍的 GA 等流量分析工具，但检查某些问题，如访问错误、蜘蛛爬行情况等，非依靠原始日志不可。大中型网站日志文件可能很大，完整人工查看是不可能的，SEO 部门可以开发专用日志分析工具。

7.4.2 常用流量分析工具

除非要查看非常细微或者只有日志中才能发现的东西，大部分情况下 SEO 人员并不会去看原始日志文件。通常网站流量分析要借助软件。

流量统计和分析软件一般分为两种。第一种是基于在页面上插入统计代码。站长在需要统计的所有网页上（通常是整个网站所有页面）插进一段统计代码，一般是 JavaScript 代码，这段代码会自动检测访问信息，并把信息写入流量分析软件服务商数据库中。统计、分析软件运行在服务商的服务器端，站长在服务商提供的线上界面查看网站流量统计和分析。

属于这类最值得推荐的是 Google Analytics（谷歌分析，GA），后面的讨论主要以 Google Analytics 为例。Google Analytics 前身是 Urchin，是一款付费统计服务。2005 年 4 月 Urchin 公司被 Google 收购，然后 Google 基于自己一贯的做法，将 Urchin 改名为 Google Analytics，免费提供给站长使用。

Google Analytics 注册地址是：http://www.google.com/analytics/zh-CN/

站长只要把 GA 统计代码放在每个页面的 HTML 中，比如全站所有页面都会调用的页脚文件中，用户访问包含有统计代码的页面时，Google Analytics 就开始收集数据。

另一个著名的服务是 Omniture ，功能强大，但价格昂贵。

类似的服务还有如下几种。

- 百度统计：http://tongji.baidu.com/
- 雅虎统计：http://tongji.cn.yahoo.com/
- 太极链：http://www.textclick.com/
- 51yes：http://count.51yes.com/
- 我要啦：http://www.51.la/
- Itsun：http://www.itsun.com/
- 阿江统计：http://www.ajiang.net

第二种是基于对原始日志文件进行分析。这类软件把日志文件作为输入，直接统计其中信息。这种统计软件既有装在服务器上的，也有运行在自己计算机上的。

装在服务器上的常见的如下。

- Webalizer：http://www.mrunix.net/webalizer/
- AWStats：http://awstats.sourceforge.net/
- Analog：http://www.analog.cx/

可以运行在自己计算机桌面的如 Azure Web Log analyzer：

```
http://www.azuredesktop.com/logdoc.html
```

这两种软件各有优缺点，比如：

- 以统计代码为基础的流量分析软件或服务的优势是简单易用，站长无须安装运行任何软件，而且软件升级维护、新功能开发都无须站长操心。
- 有的浏览器不支持（或用户有意关闭）JS 脚本，使统计代码方式数据不准。
- 有时网速慢，JS 代码又在 HTML 底部，还没来得及被运行，用户已经离开网站，也使数据不准确。
- 用户计算机或路由器或 ISP 网络上都可能存在页面缓存，用户访问页面，却可能没有在日志中留下记录。

7.4.3 流量统计分析基础

网站流量分析是个宝藏，它不仅能告诉站长网络营销的结果是什么，还能揭示原因在哪里。流量分析也是个很深的学问，涉及内容很广泛，限于本人水平和篇幅，不能在本书中详细介绍。建议感兴趣的读者参考下面这两个博客。

- 宋星的“网站分析在中国”：http://www.chinawebanalytics.cn/——深入分析又通俗易懂。
- Google 员工、GA 专家 Avinash Kaushik 的博客：http://www.kaushik.net/avinash/——最好的流量分析博客，可惜是英文的，而且是比较晦涩难懂的英文。

本节以笔者的博客为例，以 Google Analytics 为统计工具，非常简单地介绍与 SEO 最相关的几个流量分析指标，如图 7-8 所示。

图 7-8　GA 控制台

1．访问数 Visits

访问数指的是某段时间内网站被访问的总人次。如果某个用户今天访问网站，明天又来访问，将被计算为两个访问数。用户在网站上没有任何活动超过 30 分钟（GA 的定义，其他服务可能定义不同时间），也将被重新计算为新的访问数。

访问数无疑是网站流量最重要的指标之一，体现了网站推广的总体效果。流量分析软件都可以按时间，比如每天或每星期，显示出访问数。很多软件还可以以图表方式显示，更加直观。

在进行了某项营销活动后，检验效果如何的第一个指标就是看所带来的访问数。比如网站被名人在微博推荐或哪怕提到一下，经常会带来访问数的急剧提高，但通常在一两天内又会下降到和以前差不多的水平。通过访问数的变化及趋势，就可以看出营销活动的大致效果。

访问数的变化有时与营销活动无关。如图 7-8 所示，每到周末访问量都会降低，很多网站都是如此。春节期间中文网站流量下降更大。有些网站的主题是与季节或时间相关的，呈现流量起伏十分正常，不一定说明营销活动的成功或失败。

2．绝对唯一访问者数 Absolute Unique Visitors

绝对唯一访问者数指的是在某段时间内访问网站的实际人数。这个数字通常低于访问数，因为有些人会多次访问同一个网站，虽然访问数被计算为两次、三次，但只计算为一个绝对唯一访问者，如图 7-9 所示。

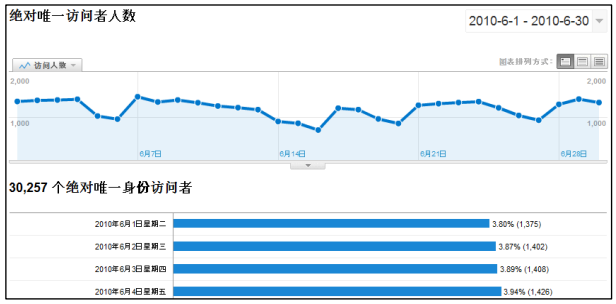


图 7-9　绝对唯一访问者数

GA 这种基于统计代码的软件一般是通过设置在用户计算机中的 cookie 识别不同用户的，所以一台电脑对应一个实际访问人数，不管他访问了多少次。但也不能排除一台电脑有多个人使用、用户删除 cookie、一台计算机使用多个浏览器等情况，这些都将使绝对唯一访问者数不会 100%准确。

基于日志文件的统计服务一般是以 IP 地址识别不同用户的，一个 IP 地址对应一个用户。通常我们说网站流量是多少个独立 IP，指的就是这个数字。但独立 IP 也不能与独立用户完全对等，比如公司局域网内的所有用户可能只记录下一个共用 IP 地址。所以没有方法能使绝对唯一访问者数 100%准确，基本准确和了解变化趋势对 SEO 人员就足够了。

3．页面访问数 Pageviews

页面访问数指的是在某段时间内被访问的页面总数，如图 7-10 所示。GA 中称为“所有访问者的综合浏览量”。这就是站长谈论流量时经常提到的 PV，英文 Pageviews 的缩写。PV 是网站卖网络显示广告时的重要依据。

通常用户访问网站时会访问不止一个页面，所以页面访问数比访问数高。

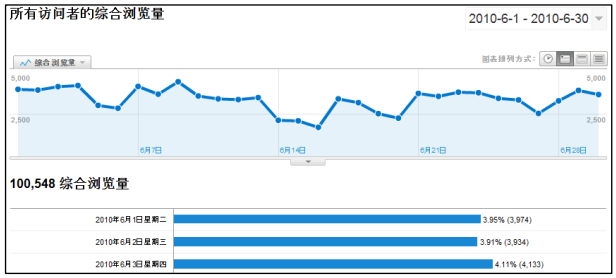


图 7-10　页面访问数

4．平均页面访问数 Average Pageviews

就是页面访问数除以访问数，Pageviews/Visits，也就是用户访问网站时平均看了多少个页面。如图 7-11 所示，“SEO 每天一贴”2010 年 6 月份平均页面访问数是 2.52，也就是说浏览者来到博客后，平均看了 2.52 个网页，如图 7-11 所示。

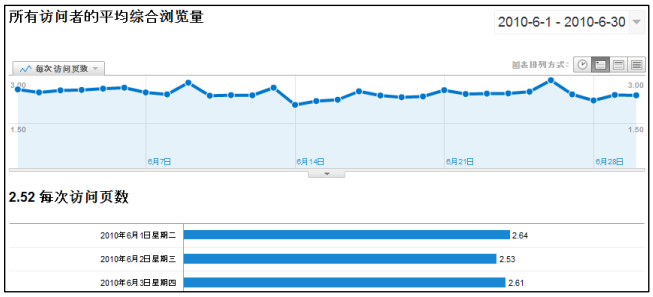


图 7-11　平均页面访问数

平均页面访问数代表了网站的黏度。一般而言，网站质量越高，黏度越高，用户看的页面越多，平均页面访问数也就越高。

平均页面访问数和网站类型也直接相关。最典型的情况是，论坛、社区类网站通常黏度很高，平均页面访问率也比较高，常常达到十几页以上。

不同流量来源的平均页面访问数也不同，说明不同流量来源的质量有区别。在后面的流量来源部分还会提到。

5．跳出率 Bounce Rate

浏览者来到网站，只看了一个网页，没有点击页面上的任何链接就离开，称为跳出。这些只查看了一个页面的访问占总访问数之比就是跳出率，如图 7-12 所示。

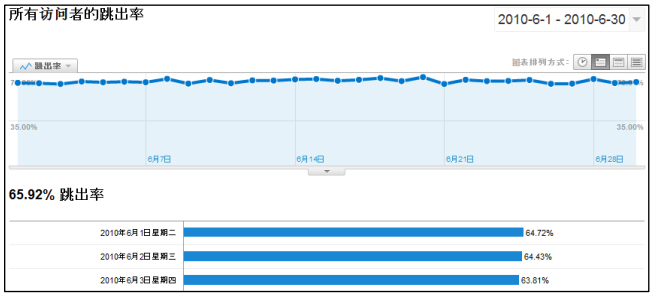


图 7-12　跳出率

跳出率是判断网站是否满足用户需求的重要指标。如果大部分用户来到网站，只打开一个网页，再也没有点击其他链接、看其他网页就离开，很可能说明网站易用性很差，用户不知道下一步该做什么，或者内容很不相关，无法吸引用户继续看其他页面。

如图 7-12 所示，我博客的平均跳出率是 65.92%，这是一个比较高的跳出率，原因

在于大部分流量是博客读者，而大部分读者是来看最新文章的，看了以后就离开，所以博客跳出率偏高是普遍现象。

如果一个普通信息类网站或电子商务网站跳出率达到 60%左右，是一个非常值得注意的警讯，应该研究到底哪些页面跳出率高，为什么如此。

6．流量来源 Traffic Sources

所有的流量分析软件都会清楚地显示三种主要流量来源所占的比例，各自的流量细节以及随时间的变化趋势。这三种主要流量来源是：直接访问、点击流量、搜索流量，如图 7-13 所示。

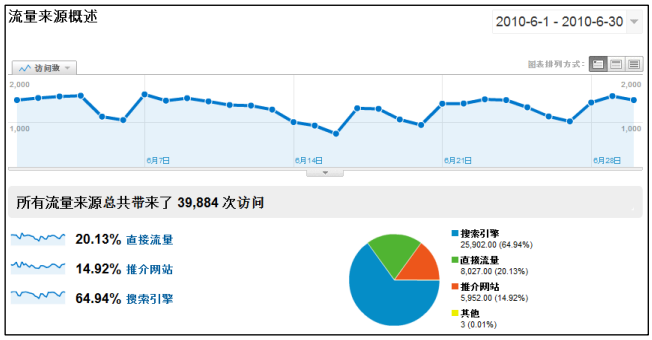


图 7-13　流量来源

直接访问，direct traffic，指的是用户通过存在浏览器里的书签或直接在浏览器地址框打入网址等直接方式访问网站。直接流量在一定程度上代表了网站忠诚用户的多少，因为只有用户觉得网站对他有帮助，才有可能存入书签或记住域名。

第二部分是来自其他网站的点击流量，referring sites，GA 中称为“推介网站”，也就是说网站链接出现在其他网站，用户点击了链接后访问网站。点击流量可能是其他网站、博客、论坛等有人提到你的网站，也可能是站长自己在其他网站购买的网络广告、交换链接等。

第三部分是 SEO 最关心的搜索流量，search engine traffic，也就是用户在百度、Google 等搜索引擎搜索关键词后，点击搜索结果访问网站。搜索流量代表了网站的 SEO 水平。

从图 7-13 可以看出，“SEO 每天一贴”直接流量比例是 20.13%，点击流量 14.92%，搜索流量 64.94%。点击流量比例偏低，搜索流量偏高，这一方面说明博客搜索引擎排名表现还不错，同时也说明流量太过于依赖搜索引擎。纯商业网站搜索流量占这么高的比例就有一定的危险性，搜索引擎算法的微小改变都可能对一个严重依赖搜索流量的商业网站造成致命打击。

流量分析软件除了显示三种流量来源各占比例，还显示流量质量细节，包括跳出率、平均停留时间、平均页面访问数等，如图 7-14 所示。

如图 7-14 所示，来自百度贴吧的流量质量比较高。如果要找友情链接，或者其他合作方式，带来高质量流量的网站是首选，内容相关，用户兴趣吻合，还能带来流量。

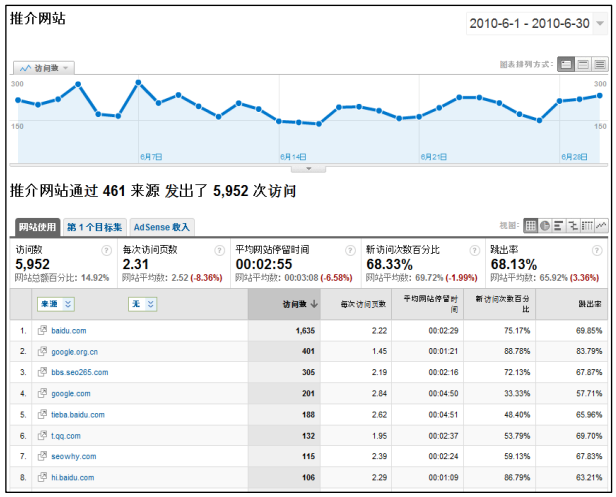


图 7-14　点击流量细节

这里所说的流量质量不一定代表推介网站本身的质量，更多地是指带来的流量相关性高低。有的时候某个网站本身质量非常高，但由于链接文字的误导性，或者设计方面的误导性，用户很容易点击链接来到另一个网站，实际上有一部分用户是误操作。这样的流量对目标网站来说相关性将大大降低，跳出率一定升高。

图 7-15 所示的关键词流量细节更是 SEO 关注的内容，流量分析中最关注的大概就是这部分。

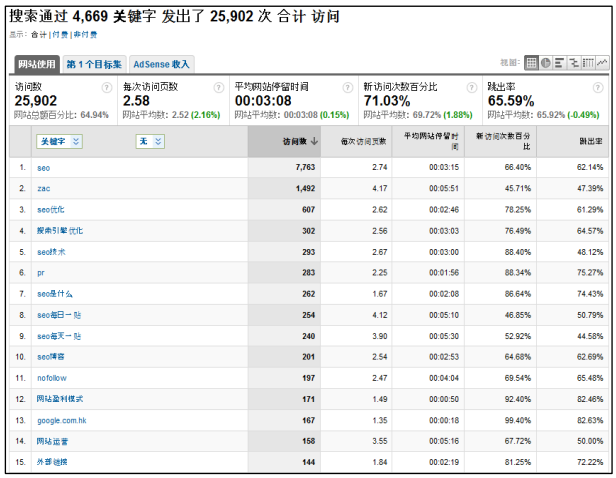


图 7-15　搜索流量按关键词细分

针对关键词的统计数字清楚显示出：

- 哪些关键词带来的流量最大？是否符合预期？

- 哪些关键词带来的流量黏度高，转化率高，对网站价值更大？
- 哪类主题的内容关键词多？对网站搜索流量贡献更大？
- 哪类关键词对实质收入贡献更大？
- 应该在哪类内容上花更多的时间？
- 哪些没有刻意优化的关键词反倒带来质量很高的流量？
- 同一个关键词在不同搜索引擎表现是否不同？怎样改进？
- 什么关键词是自己以前没有想到的却有了排名和流量？能否扩展？

……

翻看流量关键词数据既是一个 SEO 人员必做的工作，很多时候也是很有意思的工作，经常会有意想不到的收获和乐趣。

7．最常见进入页面 Top Landing Pages

指的是用户来到网站时首先访问的那个页面，或者说进入网站的入口页面。也可以称为着陆页面（PPC 领域通常这样称呼），GA 中称为目标页面，如图 7-16 所示。

图 7-16　最常见进入页面

用户进入网站，虽然很大一部分是从首页开始的，但也有一部分用户是从栏目页或具体内容页进入的，尤其是大中型、长尾做得好的网站。最热门进入页面通常就是网站上最吸引用户、最受欢迎的内容，流量可能是来自搜索引擎，也可能是其他网站觉得你的文章好而链接过来的。

SEO 人员可以检查排在前面的进入页面是不是计划中专门优化的重点页面？表现好的页面有什么共通之处可以应用到其他页面？进入页面的关键词是不是页面的目标关键词？

8．转化 Goal Conversion

排名、流量都是手段，不是目的，转化才是网站存在的目的。如 7.2 节所讨论的，

确定合适的网站目标后，还需要记录每一次转化，统计转化数据，改进网站易用性、销售流程等，提高转化率。Google Analytics 等流量分析服务可以完成简单的转化跟踪、统计，设定一个网站目标完成页面，用户访问了这个特定页面，就计算为一个转化，如图 7-17 所示。

如电子商务网站，每当有用户来到付款确认完成页面，流量分析系统都会记录网站目标达成一次。以订阅电子杂志为网站目标时，每当有用户访问订阅确认页面或感谢页面，流量系统就记录网站目标达成一次。

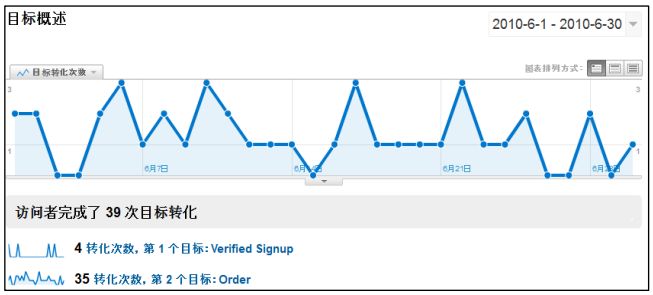


图 7-17　网站目标转化

GA 还可以列出不同流量来源、不同关键词的转化率，这是非常重要的统计数字。前面提到的黏度指标如跳出率、网站停留时间、平均页面访问数等，固然在一定程度上说明了流量质量，但是都没有最终的转化数据更能表明商业价值。高转化率的流量来源或关键词不一定就是黏度最高的来源或关键词，比如一部分忠实用户经常访问网站，却不一定有购买行为。

转化数据结合访问数、平均每单销售额和利润，可以较为准确地计算出各种流量来源、关键词的价值，这就为进一步扩展 SEO、购买广告及 PPC 定价等奠定了基础。

7.5　策略改进

研究记录 SEO 数据，除了能验证 SEO 效果、为其他部门提供数据外，更重要的是能发现问题、改进 SEO 策略。不同网站可能遇到的情况和问题千差万别，通过效果监测发现问题并没有一定的套路，SEO 人员必须深入研究数据，积累经验。下面列举一些常见到的情况。

7.5.1　收录是否充分

最简单的是查询总收录数，与网站实际页面总数对比。但是发现收录不充分，不一定能发现问题所在。通常 SEO 人员还需要研究到底哪个部分收录不充分？还是整个网站所有分类收录都有问题？如果所有分类收录比例都很低，可能是域名权重过低，网站导航出现技术问题，或者存在大量复制内容。

大中型网站经常遇到的另一种情况是，某些分类收录不错，另外一些则有问题。发现到底哪些分类收录不充分，才能检查是否链接结构有问题，指向这些分类的内部链接太少？是否需要给收录不充分的分类建立外部链接？有问题的分类内容和页面优化上有什么与其他分类不同的地方？

7.3.1 节已经讨论了 Title 和 URL 比较规则时的收录查询，再举一个页面标题和 URL 都不规则的实际例子。本书案例网站经过 URL 静态化及内部链接调整后，收录大幅上升。不过我发现一个奇怪的问题，产品列表的翻页页面收录很不均匀。比如 Title 中带有“第 2 页”，也就是第二个翻页收录只有 500 多个，这与实际存在的有翻页的分类个数相差很大，如图 7-18 所示。

图 7-18　第二个翻页收录 500 多

一直到第 8 页，收录数都是几百个，如图 7-19 所示。

图 7-19　第 8 个翻页收录 100 多

而从第 9 页开始，收录数上升到 9000 多，如图 7-20 所示。

图 7-20　第 9 个翻页收录 9000 多

后面的翻页被收录数字都维持在数千个页面以上，一直到第 50 页都还有 2000 多，如图 7-21 所示。

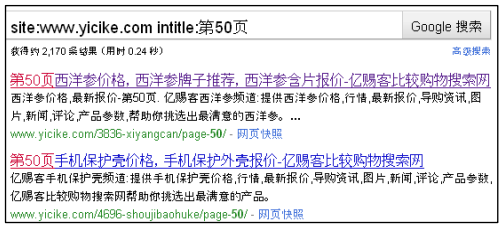


图 7-21　第 50 个翻页收录 2000 多

这说明翻页链接第 2～8 页出现了没有预见到的问题，有待研究和解决。

7.5.2　哪些页面带来搜索流量

GA 网站内容—最常见目标网页部分显示用户从哪些页面进入网站，如图 7-22 所示。

图 7-22　用户从哪些页面进入网站

选择“搜索流量”群体细分，可以知道哪些页面带来了搜索流量。点击其中任何一个页面，GA 显示出用户进入页面的来源（包括直接点击、搜索引擎及点击流量），还列出其搜索流量是由哪些关键词带来的，如图 7-23 所示。

SEO 人员可以清楚地看到哪些页面产生了搜索流量，为什么会有搜索流量，用户搜索的是什么关键词。这在一定程度上表明，在搜索引擎眼里哪些页面更加重要，与 SEO 人员计划中的重要页面是否相同，哪些重要分类或热门产品页面需要加强。

图 7-23　进入页面详情

7.5.3 目标 URL 排名如何

在关键词研究一章提到过，优化网站时需要把重要关键词分布到相应页面上，这是一个主动的过程，不能让搜索引擎自己去判断。这些被分配了关键词的页面到底表现如何，通过流量分析也可以看出。

比如我的博客首页以“SEO”为目标关键词，“SEO 技术”为分类页面：

http://www.chinamyhosting.com/seoblog/category/seo-tips/

以“SEO 技术”为目标关键词。GA 数据显示，这个分类页面在 Google 上表现不错，在百度表现相对要差很多，如图 7-24 所示。

图 7-24 “SEO 技术”分类页面流量来源

“SEO 技术”算是 SEO 领域中比较重要的关键词，这个页面在百度表现不好，其实可以重新考虑策略，把“SEO 技术”这个词分配到博客首页。

再查看如图 7-25 所示页面的关键词。

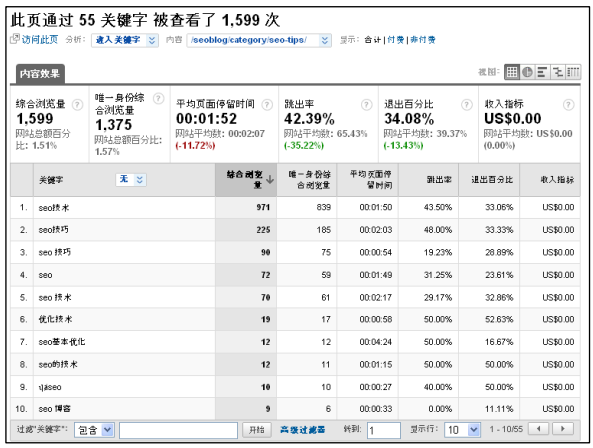


图 7-25 “SEO 技术”分类页面关键词

用户确实是搜索“SEO 技术”来到这个页面，符合预期和计划。同时发现还有用户是搜索“SEO 技巧”和“SEO 基本优化”等词来到这个页面。“SEO 技巧”是我以前没有想到的，如果真的想充分优化这个博客，可以把“SEO 技巧”这个词刻意融入到这个目标页面，而不是靠页面上偶然出现的“技巧”一词。

7.5.4 挖掘关键词

在关键词研究一章也提到过，流量分析是发现新关键词的途径之一。在 GA 中查看带来流量的关键词，除了预期中的目标关键词，还经常发现自己以前完全没想到的、五花八门的搜索词。其中与网站主题不相关，只是因为页面上偶然提到而来的流量可以不用管它。对那些与网站主题有一定相关度、以前却没发现的关键词，应该考虑是否可以增加这方面内容。把这些关键词输入 Google 关键词研究工具，看看还能生出哪些相关关键词，如果搜索量足够大，甚至可以考虑增加栏目或专题。

在关键词列表中也经常可以发现限定词。如果某些限定词表现出一定的规律性，带来不少流量，可以考虑将这些限定词优化进相应页面。

查看带来流量的关键词时，另一个要关注的是转化率。不同关键词的转化率差别相当大，常常看不出有明显逻辑性，但事实如此，如图 7-26 所示。

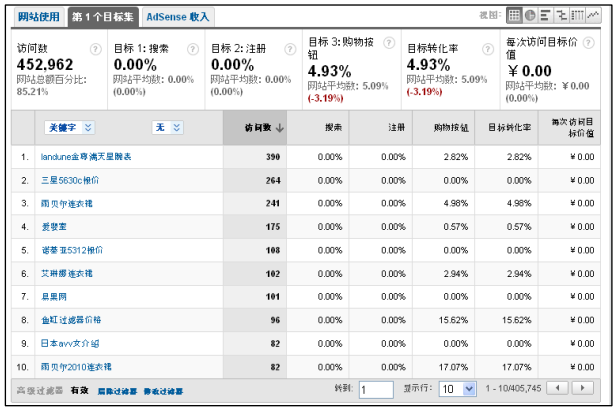


图 7-26 不同关键词转化率

找到转化率高的关键词，就找到了应该投入更多链接、增加更多相关内容的关键词。有时候，转化率不高的关键词是因为对应的进入页面内容相关度不够，需要改进页面内容，或者将关键词重新分配到更相关的页面上。

7.5.5 其他搜索引擎流量

做中文网站大家都关注百度和 Google，做英文网站一般只关注 Google，而忽视了其他搜索引擎，如雅虎、必应、搜狗等。其实这些第二集团的搜索引擎有时也能带来不少流量，尤其是英文网站。

随着雅虎舍弃自己的搜索技术，转用微软必应数据，必应在搜索引擎市场也将占据相当大份额。

从 GA 列出的搜索流量来源可以看到，虽然雅虎、必应带来的流量不多，但它们的流量质量比较高，尤其是必应。如果网站在必应排名和流量表现太差，SEO 人员可能需要投入更多精力。在相同流量情况下，必应带来的转化及赢利可能远远高于百度和 Google，如图 7-27 所示。

搜索通过 11 来源 发出了 454,007 次 合计 访问

显示：合计 | 付费 | 非付费

网站使用　第 1 个目标集　AdSense 收入　视图：

访问数	目标 1: 搜索	目标 2: 注册	目标 3: 购物按钮	目标转化率	每次访问目标价值
454,007 网站总额百分比: 85.41%	0.00% 网站平均数: 0.00% (0.00%)	0.00% 网站平均数: 0.00% (0.00%)	4.95% 网站平均数: 5.09% (-2.68%)	4.95% 网站平均数: 5.09% (-2.68%)	￥0.00 网站平均数: ￥0.00 (0.00%)

	来源	访问数	搜索	注册	购物按钮	目标转化率	每次访问目标价值
1.	baidu	349,457	0.00%	0.00%	4.57%	4.57%	￥0.00
2.	google	100,877	0.00%	0.00%	6.19%	6.19%	￥0.00
3.	images.google	2,304	0.00%	0.00%	7.90%	7.90%	￥0.00
4.	yahoo	1,088	0.00%	0.00%	6.34%	6.34%	￥0.00
5.	bing	167	0.00%	0.00%	13.17%	13.17%	￥0.00
6.	ask	49	0.00%	0.00%	6.12%	6.12%	￥0.00
7.	aol	23	0.00%	0.00%	13.04%	13.04%	￥0.00
8.	search	21	0.00%	0.00%	0.00%	0.00%	￥0.00
9.	msn	10	0.00%	0.00%	40.00%	40.00%	￥0.00
10.	yam	10	0.00%	0.00%	0.00%	0.00%	￥0.00

过滤"来源": 包含　开始　高级过滤器　转到: 1　显示行: 10　1 - 10/11

图 7-27　不同搜索引擎流量的转化率

图 7-28 是我的博客来自必应的流量。

bing 通过49 关键字 发出了 81 次 合计 访问

显示：合计 | 付费 | 非付费

网站使用　第 1 个目标集　AdSense 收入　视图：

访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
81 网站总额百分比: 0.19%	2.11 网站平均数: 2.54 (-16.88%)	00:01:27 网站平均数: 00:03:16 (-55.46%)	87.65% 网站平均数: 69.21% (26.64%)	77.78% 网站平均数: 65.43% (18.87%)

	关键字	访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
1.	google pr	10	1.00	00:00:00	90.00%	100.00%
2.	seo技术	7	1.86	00:00:52	71.43%	85.71%
3.	seo	5	1.20	00:00:04	80.00%	80.00%
4.	seo是什么	5	1.00	00:00:00	100.00%	100.00%
5.	seo优化	4	1.75	00:00:24	100.00%	75.00%
6.	google pr	2	1.00	00:00:00	100.00%	100.00%
7.	seo每一贴	2	12.50	00:13:29	50.00%	0.00%
8.	seo每日一贴	2	7.50	00:23:31	50.00%	0.00%
9.	url	2	5.50	00:03:57	100.00%	50.00%
10.	机器审查 人工	2	1.00	00:00:00	50.00%	100.00%

图 7-28　必应搜索流量

显然，主关键词“SEO”排名不好。如果今后中文搜索市场份额有变化，必应的优化是必须加强的方面。

7.5.6 长尾效果

对大中型网站来说，长尾流量往往占总流量的很大一部分。而网站是否达到应该有的长尾效果，靠检查排名很难确认，通常要从流量分析入手。

图 7-29 是我的博客搜索流量关键词分布。

搜索通过 4,854 关键字 发出了 27,032 次 合计 访问

显示：合计 | 付费 | 非付费

网站使用 | 第 1 个目标集 | AdSense 收入

视图：

访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
27,032 网站总额百分比：64.65%	2.55 网站平均数：2.54 (0.46%)	00:03:12 网站平均数：00:03:16 (-1.91%)	70.88% 网站平均数：69.21% (2.40%)	65.58% 网站平均数：65.43% (0.22%)

	关键字	访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
1.	seo	7,681	2.69	00:03:20	64.47%	61.46%
2.	zac	1,442	4.02	00:05:35	44.66%	50.28%
3.	seo优化	764	2.62	00:02:55	76.96%	59.82%
4.	seo技术	357	3.34	00:04:50	84.59%	43.70%
5.	搜索引擎优化	318	2.35	00:02:53	73.27%	68.24%
6.	pr	311	1.87	00:01:30	87.46%	81.67%
7.	seo是什么	311	1.78	00:02:19	88.75%	76.21%
8.	拜里巴巴网站	277	1.16	00:00:15	98.19%	90.61%
9.	seo每天一贴	262	3.98	00:06:31	47.71%	38.55%
10.	seo每日一贴	246	3.72	00:04:41	46.75%	49.59%
11.	seo博客	223	2.70	00:03:28	63.23%	60.54%
12.	网站运营	215	4.21	00:06:32	70.70%	45.12%
13.	网站盈利模式	184	1.48	00:00:45	92.93%	82.07%

图 7-29　带来搜索流量的关键词

来自“SEO”和“ZAC”这两个词的流量就占了三分之一。前十个关键词带来的流量占总搜索流量的 50%。这说明我的博客远远没有发挥长尾效应。当然原因主要在于博客规模太小，只有几百篇帖子。对大中型网站，尤其是收录几十万、上百万的网站来说，如果前 10 个关键词占到流量的 50%，说明网站内部链接权重分配和页面基本优化有重大问题，没有发挥出大型网站应有的潜力。

7.5.7 关键词排名下降

从连续记录的关键词排名数据应该可以看到哪些关键词排名在下降。很多时候一部分关键词排名下降，另一部分上升很正常，尤其是下降幅度在几位或十几位之内时。但如果大部分关键词排名同时下降，很可能说明网站被惩罚，或者搜索引擎算法的某个改动影响整体网站排名。

遇到关键词排名整体下降时，通常建议先按兵不动一个月左右。有时候排名下降是因为搜索引擎算法改变引起的，搜索引擎还可能会继续微调算法，最终使排名恢复。

如果一到两个月后排名没有变化甚至下降更多，则需要研究下降后的位置被哪些网页占据，排在前面的有没有新出现的竞争对手，这些页面有什么特点，有没有共性，有哪些方面是与自己网站不同的。

7.5.8 链接诱饵成效

一个成功的链接诱饵到底吸引了多少外部链接很难确切知道，使用 Yahoo! Site Explorer 外链工具也不容易衡量。Majestic SEO 是这方面最好的工具。输入域名后，Majestic SEO 用曲线形式清楚地显示出在哪段时间外部链接有大量增长。当然，Majestic SEO 显示的曲线也表现出外部链接整体工作的成效。

7.5.9 发现链接伙伴

在 GA 中查看来自其他网站的点击流量，也是个很有意思、有时有意想不到回报的工作，因为经常能看到自己从没听说过的网站，之所以会有点击流量过来，一定是对方链接到自己的网站，如图 7-30 所示。

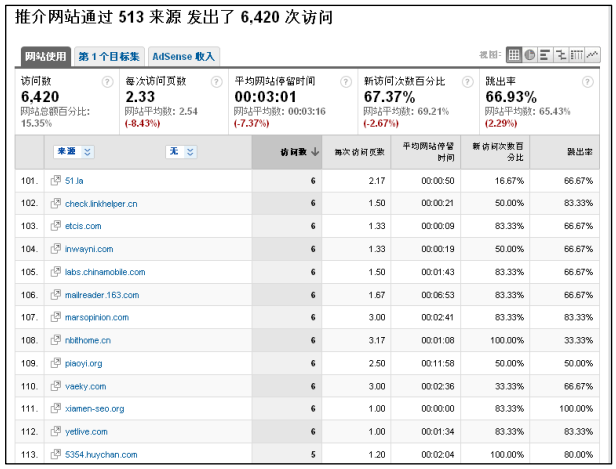


图 7-30 点击流量来源可能也是最好的链接来源

访问对方网站，看看链接出现在什么地方，对方站长为什么会链接过来，是否因为对自己网站上的某类内容很感兴趣，在评论什么内容或功能，今后增加这方面内容，是否能够吸引更多类似链接。有的时候也可以主动通过某种方式，如博客留言等，与对方站长增强联系，制造更多的链接机会。

7.5.10 寻找有潜力关键词

Google 网管工具现在列出网站在不同关键词下的平均排名，如图 7-31 所示。SEO 人员可以看看有哪些关键词能够通过最少的工作，获得最多流量增长。

比如平均排名处于第 2 页最前面的关键词，只要稍微调整相应页面，或者增加几个外部链接，就有可能使这些页面进入第 1 页。我们从搜索结果点击率分布知道，第 2 页上的结果获得的点击远远少于第 1 页，而第 1 页 6～10 名获得的点击，又远远少于排在

前三位的结果。第 2 页最前面和第 1 页 6～10 名就属于投资回报率比较高的，只要花不太多的精力，就有可能提高几位排名，获得数倍的流量。而排在第七八页的页面，即使通过努力爬到第三四页，也不会带来明显流量增加。

查询	展示	点击次数	点击率	平均排名
seo优化	3,600	260	7%	2.6
seo是什么	2,900	320	11%	3.6
网站关键词优化	1,900	<10	-	3.4
zac	1,600	260	16%	9.2
搜索引擎优化	1,300	110	8%	2.4
gg	1,300	<10	-	29
google.hk	1,000	<10	-	9.8
搜索引擎排名	880	<10	-	9.8
seo	720	36	5%	70
每天	720	<10	-	4.7
google 退出中国	720	16	2%	11
google.cn	720	<10	-	9.5
seo技术	720	260	36%	2.5
虚拟主机	720	<10	-	51
网络营销	590	22	4%	77
创业点子	590	22	4%	16
google.com.hk	590	12	2%	10
邮件营销	480	<10	-	18
实时搜索	480	73	15%	3.1
杆具	480	<10	-	12
google站长工具	480	28	6%	3.9
网页打开慢	390	22	6%	5.4

图 7-31　Google 网管工具显示的关键词排名

第 8 章

SEO 作弊及惩罚

本书主要介绍正规 SEO 手法。本章简单介绍一些 SEO 作弊手法，以及可能带来的搜索引擎惩罚。

在 SEO 行业通常把作弊称为黑帽 SEO，把正规手法称为白帽 SEO。使用黑帽或作弊手法有时是因为站长不了解而无意中误用，也有很多时候是为了提高排名和流量而刻意使用。作弊经常导致被搜索引擎惩罚，甚至导致网站被完全删除。

8.1 白帽、黑帽、灰帽

在搜索引擎优化行业，一些不符合搜索引擎质量规范的优化手法，也就是作弊的 SEO 手法被称为黑帽，英文为 Blackhat。而正规的符合搜索引擎网站质量规范的就称为白帽，Whitehat。由于搜索引擎公布的质量规范和准则比较笼统，常常有各种解释的空间，那些不能被明确归入黑帽或白帽，介于两者之间的优化手法就被称为灰帽 SEO（Greyhat）。

8.1.1 白帽、黑帽是风险度判断

虽然被称为“作弊”，但大部分黑帽 SEO 还谈不上是一个价值判断，只是站长的一个风险度判断而已。很多黑帽手法只涉及站长在自己的网站上进行超出搜索引擎规范的操作，并不直接影响其他网站。所以就算是作弊，也谈不上道德问题。虽然主流 SEO 行业都不提倡黑帽，但也并不是从价值观角度出发的。

黑帽 SEO 往往会导致被搜索引擎惩罚，站长需要自行判断：风险到底有多大？网站作弊可能带来的流量与可能的惩罚相比到底哪个更划算？对一个商业网站来说，能承受的商业风险有多大？网站被惩罚或删除对品牌和口碑影响有多大？公司必须事先做个判断。

白帽 SEO 认为，用黑帽技术可以得到排名和流量，但是过不了多久网站就会被惩罚，被封掉，又得重新去做另外一个网站。为什么不用白帽手段踏踏实实地做一个健康的、对人们真正有用的网站呢？

白帽网站不用担心会被搜索引擎封掉，站长也可以骄傲地说：这个网站是我的。十年、二十年以后，这个网站还在给你带来流量和利润，何乐而不为呢？

但站在黑帽 SEO 的角度，他们也有他们的道理。很多黑帽 SEO 是使用程序生成网站的，他们建立一个几万、几十万页的网站不费吹灰之力，只要放他们的蜘蛛出去抓取就可以了。就算过几个月网站被惩罚，他可能已经赚了一笔了，投资报酬率还是相当高的。

白帽 SEO 所关注的是长远利益，也就是两年三年，甚至十年八年以后的利益。如果你坚持认真做站，不使用作弊手段，坚持十年，如果不出大的意外，你的网站一定会得到很好的流量。有了流量，就有了赢利。十年以后，你的网站有稳定赢利时，对搜索引擎的依赖就小多了。你也不必每天花那么长时间在你的网站上，你的网站会自动带来源源不断的流量。

站在黑帽 SEO 的立场上，这种放长线掉大鱼的策略，即使很正确，有的人也不愿意这么做。认真建设网站经常是一件很无聊的事，你要写内容，做关键词研究，做流量分析，和用户交流沟通等。更何况要坚持几年！

黑帽 SEO 要做的就简单多了。买个域名，甚至可以就使用免费子域名，连域名都省了。主机也有很多免费的。程序一打开，放上 Google AdSense 代码，到其他论坛或博客留言，这些留言也有可能是程序自动群发的，然后就等着收支票了。

黑帽 SEO 还有一个无法否认的论据是：你不能保证完全遵守搜索引擎的规则，就能在十年后得到一个受搜索引擎重视的网站。谁知道搜索引擎什么时候会对算法做一个大的改变，让成千上万白帽网站从搜索引擎里消失呢？

黑帽 SEO 赚钱的短平快方法，也有它的优势。

所以说，抛开道德观念不谈，黑帽和白帽 SEO 的选择，更多的是对风险度的判断以及对生活方式和企业模式的选择。你是要花时间和精力建立一个长久健康的企业？还是轻松简单地赚一笔快钱，但是要冒很高的随时被封杀的危险？

要做黑帽 SEO，就要在不影响其他人的前提下，自己对自己负责。

8.1.2 道德及法律底线

也有一部分黑帽手法超过了道德或法律的底线。比如在其他网站上大量群发垃圾留言，就不仅仅是在自己网站上操作，而直接影响到其他网站。轻者给其他网站带来删除垃圾的工作量，重者影响对方网站质量，这就已经属于不道德的范畴了，应该坚决反对。

再比如使用黑客手法攻入其他网站，加上自己的链接，这就已经超过了法律底线，变成了违法行为，就更是不能容忍。这是黑客，不是黑帽，更不是 SEO。

8.1.3 SEO 服务商的底线

在自己网站上使用黑帽手法是自己的事，被惩罚或删除也不能怪别人。作为 SEO 服务商在客户网站上使用黑帽手法，则需要非常小心。如果是竞争很强的行业，不尝试黑帽无论如何也没有效果时，必须事前清楚地告诉客户搜索引擎对黑帽手法的态度，以及黑帽带来的风险，由客户自行判断和取舍。只有在客户清楚风险并且能够承担后果的情

况下，才能在客户网站上尝试黑帽手法。

8.1.4 黑帽 SEO 的贡献

一些黑帽 SEO 人员对搜索引擎排名算法了解很深，技术也很高超。这类黑帽 SEO 在自己的网站上做实验，测试搜索引擎的底线，对所有 SEO 行业的从业者都是不小的贡献。没有黑帽的探索和尝试，我们往往就不能明确知道哪些手法被搜索引擎接受，哪些会触及搜索引擎的底线。在这个意义上，不逾越道德和法律界限的黑帽对整个 SEO 行业并不是一件坏事。

8.1.5 承担风险，不要抱怨

虽然说很多黑帽 SEO 手法谈不上不道德，但搜索引擎肯定是痛恨黑帽的，因为黑帽降低搜索结果相关性和用户体验，动了搜索引擎的饭碗，属于必须严厉打击的范围。所以做黑帽被惩罚是正常的，甚至可以说，长远来看，是必然的。

在了解黑帽风险的基础上，如果自己使用了黑帽被搜索引擎惩罚，就不必抱怨，只能自己承担后果。

经常在论坛上看到站长抱怨自己网站排名下降，被搜索引擎惩罚甚至删除，站长感觉很无辜，认为搜索引擎随随便便没有原因地惩罚网站。这种抱怨其实往往是误导，深入研究下去就会发现，被惩罚的原因恰恰是使用了被搜索引擎认为是作弊的手法。

举一个典型的，我觉得还挺有喜剧效果的例子。2009 年 5 月 21 号，Google 英文帮助论坛里一个站长发帖，貌似很无辜地被 Google 删除了所有网站。他列出了 5 个信息类网站，据他所说都是真正的、人工做的信息类网站，针对不同话题，没有一个是那种自动生成内容的垃圾网站。5 个域名，无一例外都是充满关键词的长域名。

站长很迷惑很委屈地问：这是为什么？这是为什么？这不公平啊！

然后有人指出，这几个网站都有很大一部分是专门用来做友情链接交换的页面，所有这些网站互相交叉链接起来，形成链接农场。站长又很纯真地回答，我们这些交叉链接不是为了 Google PR 和链接，因为我们早知道同一个 IP 上的交叉链接是没有用的。当然，这位站长不在乎 PR 和链接，但很在乎网站被删除。

又有人发现网站的一些页面有重复内容，最典型的是隐私权政策页面。这倒不是什么大问题，隐私权政策本来就都差不多。这位站长严正声明，我们公司有政策，要求员工不可以抄袭内容，所以大部分内容都是原创的。

然后一位 Google 员工忍不住跳出来指出，这位站长网站上经常有类似这样的内容：

Every year, millions of people suffer from head injury symptoms. Most of these are

minor because the skull is designed to protect the brain. Most closed head injury symptoms will usually go away on their own. However, more than half a million head injuries a year, are severe enough to require hospitalization.

而这些内容在其他网站上早就有了，只不过差几个字：

Every year, approximately two million people sustain a head injury. Most of these injuries are minor because the skull provides the brain with considerable protection — thus symptoms of minor head injuries usually resolve with time. However, more than half a million head injuries a year are severe enough to require hospitalization.

很显然，这位站长所谓的原创就是加减一些“的、地、得、可能、也许”之类的词，再换换段落和语序什么的。国内很多站长所说的伪原创也是以这种技巧为基础吧？请记住，“伪原创”就是说，不是原创。

然后这个站长又不经意地提到，他的一个合作伙伴的哥哥，针对不同关键词做了很多所谓原创内容网站，问题在于，数量多达好几百个，还真是惊喜连连。

有人列出来一些这位站长的网站，点进去看看就知道是典型的单页面站群，只有一个页面，两三段文字，而且还是抄来的，然后放上 Google AdSense 或联署计划链接，整个网站剩下的就是交换链接页面。刨去一页抄袭内容、广告和交换链接，就什么都没有了。

这位站长还是觉得很无辜，百思不得其解，提出一个最可爱的问题：有没有可能 Google 的工程师能跟我们当面谈一下，告诉我们原因是什么呢？不用 Google 工程师到我们这里来，我们愿意去 Google 拜访……

我也尝试过查看几个站长抱怨被无故惩罚的中文网站，无一例外，要么有作弊嫌疑，要么是无意义的内容抄袭网站，本不该有好的排名。

8.1.6 了解黑帽，做好白帽

在这里介绍黑帽，并且说黑帽不一定是不道德的事，并不意味着鼓励大家使用黑帽。恰恰相反，对一个正常的商业网站和大部分个人网站来说，做好内容，正常优化，关注用户体验，才是通往成功之路。

要做好白帽也必须了解黑帽都包括哪些手法，避免无意中就使用了黑帽手法。我自己在刚开始接触 SEO 时，就曾经天真地使用与背景颜色相同的隐藏文字，当时丝毫不觉得有什么不妥，完全是无意中使用的。

有些黑帽手法风险更大，近乎于杀无赦。只要被发现，网站一定会被惩罚或删除，比如隐藏文字。有的黑帽手法风险要低一点，搜索引擎还会考虑网站的其他因素，惩罚比较轻微，也有恢复的可能，比如关键词堆积。

做白帽就要花更多时间和精力，而且并不能百分之百保证一定能做出一个成功的网站。但相对来说白帽也更安全，一旦成功，网站就可以维持排名和流量，成为一份高质量资产。

黑帽手法常常见效快，实施成本低。问题在于被发现和惩罚的概率很高，而且会越来越高。一旦被惩罚，整个网站常常就不得不放弃了，一切要重新开始。长久下去，很可能做了几年后，手中还是没有一个真正高质量、能被称为资产的网站。

8.2 主要 SEO 作弊方法

8.2.1 隐藏文字（Hidden Text）

隐藏文字指的是页面上出现用户看不到，但搜索引擎能看到的文字。当然，这些文字是以搜索引擎排名为目的，所以通常包含大量关键词，意图提高文字相关性。有时隐藏的文字还与可见页面内容无关，目的是希望某些与本页面无关但搜索次数高的关键词能有排名和流量。

实现隐藏文字的方法有多种。比如，文字与背景颜色相同，页面背景设置为白色，文字也设置为白色。这是最简单的方法，也是最容易被检测出来的方法。

更改进一点，有的作弊者把文字放在一个图片背景上，而图片就是一个单色图案，与文字是同一个颜色，比如白色文字放在白色图片上。由于搜索引擎还不能准确读取图片内容，就可能无法判断这是隐藏文字。

有的时候作弊者也使用相近颜色。比如，背景颜色是纯白色，HTML 代码是#FFFFFF，文字设置为非常非常浅的灰色，比如使用#FFFFFD。这样，搜索引擎读到的颜色代码是不同的，但用户肉眼其实分辨不出这两种颜色的区别。

使用微小文字也可以实现隐藏文字，比如把文字大小设置为一个像素，这样用户在页面上是看不到这些文字的。

还有的黑帽使用样式表隐藏文字，近几年这种方式更流行。通过 CSS 文件把文字定位到不可见区域，比如文字放在屏幕左边或右边很远的地方：

```
positon:absolute；

margin-right:-100000px；
```

这样用户在正常情况下是看不到这些文字的。或者把文字放在不显示的层上：

```
<div style="display: none">隐藏文字</div>
```

用户正常情况下很难发现隐藏文字，但可以查看 HTML 源代码，有的隐藏文字可以

通过按 Ctrl+A 键选择页面上所有文字的方法看到。对搜索引擎来说，有些隐藏文字确实很难通过程序检测出来。但是一旦用户或竞争对手举报，人工审查就很简单了。隐藏文字是一旦被发现通常会被惩罚的、风险较大的黑帽手法。

8.2.2 隐藏链接（Hidden Links）

与隐藏文字类似，隐藏链接就是用户看不到，但搜索引擎能看到的链接。实现方法与隐藏文字也类似。

隐藏链接可能是站长在自己网站上链接到自己的页面。更常见的是黑入其他网站，加上指向自己网站的隐藏链接。这样，被黑的网站站长看不到链接，被发现和删除的可能性降低。

隐藏链接属于明确的作弊，但是在判断谁应该被惩罚上有一定难度。假设 A 网站上有隐藏链接连到 B 网站，那么搜索引擎是该惩罚 A 网站还是 B 网站呢？如果因为 A 网站上有隐藏链接而惩罚 A 网站，但其实可能是 B 网站黑进了 A 网站加上的链接，A 网站本身是无辜的。如果因为隐藏链接惩罚 B 网站，但也可能是 A 网站甚至是第三方恶意陷害 B 网站。

这时候对搜索引擎来说，比较保险的方法是使隐藏链接效果归零，另外配合网站上出现的其他作弊模式进行判断，在有确切证据证明无论是 A 网站还是 B 网站作弊情况下，还是可能被惩罚。

8.2.3 垃圾链接（Link Spam）

所有主流搜索引擎都把外部链接当做排名的主要因素之一，而从其他网站获得自然链接又不是一件容易的事，垃圾链接就应运而生了。

垃圾链接通常是指站长为了提高排名，在其他任何可以留言的网站留下自己的链接，比如开放评论的博客、留言板、论坛，文章带有评论功能的网站，社会化媒体网站等。

与正常留言不同的是，垃圾留言常常有两个明显特征。一是留言与原本的帖子或文章主题毫无关系，只是为了留下链接而留言。我们经常看到博客里有“非常同意”、“不错”之类毫无意义的话，很大一部分是垃圾链接。

第二个特征是，留下的链接中锚文字常常是目标关键词。这也就是为什么在博客评论中经常看到把自己称为“加湿器”、“起重机”、“深圳搬家”、“注册公司”等的留言，其目的都是为了这些关键词的排名。正常的有礼貌的留言者应该留下自己的真实姓名或网名，而不是这些垃圾关键词。

很多时候垃圾留言是通过群发软件完成的，网上还有不少人在出售这种群发软件。

垃圾链接判断对搜索引擎来说并不难。使用 WordPress 博客垃圾评论过滤 Akismet 插件的人都会注意到，Akismet 插件判断垃圾留言的准确率相当高。Akismet 系统会根据留言特征判断是否垃圾留言和链接：

- 留言字数。很多垃圾留言往往只是“好文章”，“说得不错”，“顶”之类的话。这种短小而又没什么意义的文字，会增加被过滤的可能性。
- 链接数量。一段留言包含一个链接是正常的，包含几个链接，还排在一起，就很值得怀疑了。
- 是否包含常见垃圾关键词。大量使用垃圾留言和链接的经常聚集于某些特定关键词，如前面提到的“加湿器”、“起重机”、“深圳搬家”、“注册公司”之类。物以类聚，人以群分，在 SEO 中也适用。
- 链接指向的网站是否在黑名单中。反垃圾软件会收集垃圾网站及 IP 地址。一旦上了黑名单，留言就直接被过滤掉了。
- 人工过滤。博主可以在 WordPress 后台人工删除没有被自动识别出来的垃圾留言，被很多博主人工删除的网站离黑名单就不远了。

搜索引擎比反垃圾插件的数据掌握和判断力无疑更强。除了上面提到的留言本身特征，搜索引擎还可以同时看到多个网站的数据：

- 留言的时间关系。比如同一个博客几秒钟内在不同帖子出现同一个网站的留言链接，这肯定不正常。或者同一个网站的垃圾链接在几乎同一时间段出现在多个博客上。单个博客作者看不到这些信息，但搜索引擎检测到这些数据易如反掌。
- 博客帖子有多老及最新留言情况。如果博客帖子已经是两年前的，而且一年半没有新留言了，冷不丁冒出一个留言，这多少也有点可疑。
- 留言相关性。搜索引擎可以通过语义分析判断留言与博客帖子是否有一定的相关性。群发软件留下的垃圾，通常都是没什么意义，放在哪儿都看似相关，也都不相关的话。
- 多个博客留言模式。群发软件都会同时往大量博客发垃圾，而且留下的域名及留言内容都一样或相似，这种模式会引起怀疑。
- 垃圾链接出现的速度。与真正的读者留下的有意义的留言不同，群发软件都是快速留下大量链接，搜索引擎也可以检测到这一点。

垃圾链接的效果是非常值得怀疑的。有的 SEO 看到一些网站排名很好，检查外部链接后发现大部分链接是垃圾链接，所以认为垃圾链接还是很管用的。其实，造成网站排名不错的很可能是其他少量的、不易被发现的高质量链接。

通常对搜索引擎来说，最简单的处理方法是直接忽略，把垃圾链接效果归零。因为垃圾链接而惩罚网站则比较谨慎，这些垃圾留言既可能是站长自己留下的，也可能是竞争对手恶意陷害。

8.2.4 买卖链接（Paid Links）

外部链接是搜索引擎排名最重要因素之一。自然获得外部链接又非常困难，需要大量高质量内容、创意、时间、精力。链接建设是 SEO 最难的部分。

相比之下，用钱买链接就显得比较简单，直截了当。不过，为了排名而买卖链接是所有搜索引擎深恶痛绝、明确反对的。

2007 年 10 月份，Google 就曾经大张旗鼓地打击链接买卖，降低了很多参与买卖链接网站的 PR 值。

2009 年 2 月份，Google 还不惜挥刀自宫，将 Google 日本网站 PR 值从 9 降到 5。Matt Cutts 确认就是因为 Google 日本网站使用了付费评论，当然评论中包含有链接。

这是给整个 SEO 行业一个鲜明的信号，Google 对链接买卖的打击是明确而且不遗余力的。其他搜索引擎也同样如此。

当然，链接买卖并不一定就是为了排名。在搜索引擎出现之前，网站之间买卖链接本来是很正常的，那时候是作为广告的一种形式，带来的是直接点击流量，而不是为了搜索排名。所以买卖链接是否算 SEO 作弊就变得比较模糊了，搜索引擎并不能百分之百准确判断出有金钱关系的链接交易是否以提高排名为目的？

搜索引擎对链接买卖的检测算法正在快速进步中，除了对手举报导致搜索引擎人工审查之外，在算法上还可能考虑以下因素：

- 内容主题是否相关？通常买链接的目标都是高 PR 值的大网站，往往就会忽略内容相关性。同行业的网站之间买卖链接比较少见，而且同行业网站之间链接也很正常。
- 链接的突然出现和消失，也是链接买卖常见特征之一。一旦买链接的网站觉得没有效果，不再继续付费，链接也就突然之间消失了。
- 链接出现的位置。买卖的链接最常出现的位置就是页脚和左右导航栏中的赞助商或广告商链接部分。也有很多买卖链接是出现在页脚的全站链接。
- 使用知名链接交易服务。搜索引擎工程师可以很容易地注册一个账号，挖出几个参与链接买卖的网站，就能带出一大群交易网络中的网站。
- 与已知买卖链接的网站有关。通常卖链接的网站不会只卖一两个链接，而是看到有利可图就会上瘾，不停地卖链接。经常买链接的网站也同样如此，会从很多不同网站买链接。所以，确认几个网站买卖链接，从链接关系上可以挖掘出更多可能买卖链接的网站。

当然，买卖链接的判断不可能百分之百准确。不要说算法，连人工审查也不可能百分之百准确。想象一下，两个站长在线下聚会时认识了，一起吃饭喝酒时就把链接买卖的事敲定了。搜索引擎不可能有方法确认这两个网站之间的链接实际上有金钱来往。

不以排名为目的的链接买卖也应该注意不要被误判为意图操纵排名。比如应该做到下面几点：

- 链接来源和锚文字多样化。既有来自高 PR 值网站的链接，也有普通甚至低质量网站的链接。锚文字既有目标关键词，也有公司名称，甚至“点击这里”之类的文字。通常买来的链接都是来自高 PR 值的网站，锚文字是目标关键词。作弊的站长一般不会花钱从一个小网站买一个锚文字为“点击这里”的链接。
- 链接增加匀速，既不要出现不成比例的大量增长，也不要突然消失一批链接。正常的网站外部链接都应该是大致匀速增长的，除了偶尔因为链接诱饵等活动有爆发点。
- 尽量寻找不太卖广告的网站。对方网站卖的广告越多，上面的链接都被判断为买卖链接的可能性就越大。
- 链接的形式多样。链接可以出现在博客的 Blogroll 中，也可以出现在正文中。既可以是文字链接，也可以采取旗帜广告的形式。既有首页链接，也可以有内页链接。在链接来源网站上，也是既有来自首页的，也有从内页获得的。
- 尽量避免链接出现在页脚中。

总之，就算是有金钱交易，也尽量把它当做商业扩展的一部分，而不是为了排名而做的链接买卖，不可 SEO 目的太强。对正规网站的 SEO 来说，链接买卖风险比较大，应该尽量避免。

8.2.5 链接农场（Link Farm）

链接农场指的是整个网站或网站中的一部分页面，没有任何实质内容，完全是为了交换链接而存在。交换链接页面上全部是链接到其他网站，其他网站也都链接回来。这些网站之间就形成了一个链接农场，互相交叉链接。

很多时候这种链接农场是同一个公司或站长所控制的一群网站，也有的时候这些网站都是参与某个交换链接联盟的网站。

更危险的是，有时参加链接农场的网站还使用相同的软件来管理页面上的链接。参加交换链接联盟的站长申请链接、批准链接都通过软件自动实现，交换链接页面也是软件自动生成。之所以危险是因为这类软件所生成的页面格式相同，有时页面的分类都基本相似，搜索引擎很容易就能判断出整个链接农场。

链接突然大量增加，大量同质外链（IP、网站结构、设计等方面的雷同），锚文字集中，内容不相关等，都是链接农场常见的特征。

检测出链接农场或有链接农场嫌疑的，搜索引擎可能会采取以下措施：

- 这些页面的链接在计算权重时完全不被考虑
- 页面上的链接权重被降低
- 页面上的链接权重被惩罚
- 这些页面本身重要性被降低
- 这些页面本身重要性被降低，同时导出的链接重要性也被降低

8.2.6 链接向坏邻居（Bad Neighborhood）

自己的网站链接到已经被判断出作弊、并且被惩罚的网站，这个作弊和被惩罚的网站就是一个坏的邻居。

坏邻居网站通常是一个特定的域名，在极少数情况下也可能是一个 IP 地址上的很多作弊网站。

所以在正常交换链接时，要大致判断一下对方网站是否作弊？是否被惩罚？坏邻居链接到自己的网站并没有关系，因为你没办法控制其他人的网站。但你链接到坏邻居，就可能连累到自己的网站也被惩罚。

8.2.7 隐藏页面（Cloaking, Cloaked Page）

隐藏页面也有人翻译为障眼法。

隐藏页面指的是页面使用程序判断访问者是普通用户还是搜索引擎蜘蛛，如果是普通用户，程序返回一个不考虑 SEO、只是给用户看的正常页面，如果是搜索引擎蜘蛛，程序就返回一个高度优化、常常优化到语句已经没办法读的程度的页面。所以普通用户和搜索引擎看到的页面内容是两个不同版本。

搜索引擎蜘蛛根据自己抓取到的页面进行排名，由于搜索引擎还不可能真正理解文字内容，所以哪怕语句完全杂乱无章，搜索引擎也判断不出来。而用户看到的却是没有充斥关键词，看起来比较自然的文字。

要判断网站是否使用了隐藏页面，可以有几种方法。

比较简单的是，访问网站时改变浏览器的用户代理（user-agent），将自己的浏览器伪装成搜索引擎蜘蛛。SEO 工具部分介绍的 SEO Toolbar 插件可以实现这个功能。火狐浏览器有不少插件也可以更改浏览器的用户代理。

不过比较高级的隐藏页面程序不仅会检查浏览器类型及版本信息，还会检查访问来自哪个 IP 地址，只有在访问者 IP 地址是已知搜索引擎蜘蛛 IP 时，程序才返回优化的版本。

另外一个判断方法是看网页在搜索引擎中的快照。如果快照中显示的和用户在浏览器中看到的内容差别巨大，就说明使用了隐藏页面技术。

近几年也经常有作弊者使用隐藏页面骗取友情链接。普通用户访问时，显示带有所有正常友情链接的页面，而返回给搜索引擎蜘蛛的版本则删除了所有友情链接。这样，与之交换友情链接的站长以为对方放了自己的链接，但其实搜索引擎完全看不到。

搜索引擎判断隐藏页面时，与用户刚好相反。几乎所有的搜索引擎都会发出匿名蜘蛛，也就是访问时模仿普通浏览器的用户代理信息，抓取页面后与正常蜘蛛抓取的数据相比较，从而判断隐藏页面。

隐藏页面与正常的 IP 传送（IP Delivery）之间容易混淆。IP 传送指的是网站程序检查来访者的 IP 地址，然后根据 IP 地址返回不同内容。

比如很多分类广告网站就使用 IP 传送方式，来自不同城市的用户将看到针对自己所在城市生成的内容。北京的用户看到的就是北京地区分类广告，上海用户看到的就是上海地区分类广告。其中还可能包括了转向，北京用户访问首页时，会被自动转向到北京地区专用的子域名或二级目录。

IP 传送的实质也是不同用户看到不同版本。隐藏页面与 IP 传送的区别在于，隐藏页面是针对普通用户和搜索引擎蜘蛛返回不同内容，IP 传送是针对不同 IP 地址返回不同内容。只要一个来自北京 IP 地址的搜索引擎蜘蛛看到的内容，与来自北京 IP 地址的用户看到的内容是一样的，就是 IP 传送，而不是隐藏页面。如果同样是来自北京的 IP 地址，但搜索引擎蜘蛛看到的内容与用户看到的内容不一样，就是隐藏页面。

隐藏页面和隐藏文字类似，属于比较严重、明确的 SEO 作弊手法。

8.2.8 PR 劫持（PR Hijacking）

使用欺骗手段获得工具条上比较高的 PR 值显示。方法是利用跳转。

前面提到过，一般搜索引擎在处理 301 和 302 转向时，是把目标 URL 当做实际应该收录的 URL。当然也有特例，不过在大部分情况下是这样处理的。

所以如果作弊者从域名 A 做 301 或 302 跳转到域名 B，而域名 B 的 PR 值比较高，域名 A 在 PR 更新后，也会显示域名 B 的 PR 值。有人就利用这一点，把自己的域名 PR 值刻意提高。

最简单的就是先做 301 或 302 跳转到高 PR 的域名 B，等工具条 PR 更新过后，立刻取消转向，放上自己的内容。用户访问这个网站 A，看到的是一个高 PR 值，却不知道 PR 值是通过转向劫持得到的，不是这个网站的真实 PR 值，而是另外一个网站的。劫持的 PR 显示值至少维持到下一次工具条 PR 更新，一般有两三个月时间。

更隐晦一点的办法是，通过程序检测到 Google 蜘蛛，返回 301 或 302 转向，对普通访问者和其他蜘蛛都返回正常内容。这样，用户看到的是普通网站，只有 Google 会看到转向。

劫持 PR 值的作弊者的目的也很明显，就是为了卖链接、卖 PR。如果卖链接的诉求是广告性的直接点击流量，还有情可原，如果主要诉求或唯一诉求就是高 PR 值，这无疑就是欺骗，这种网站上的链接对 PR 值没有任何贡献。在寻找和买链接的时候，如果对方炫耀的就是高 PR 值，就要非常小心。

怎样鉴别劫持得来的 PR 值呢？最准确的当然就是看 Google 的网页快照，如果你看到的网页是一个样子，Google 快照显示的却是另外一个样子，网站标题和 logo 都是另一个网站，这很可能就是 PR 劫持了。

8.2.9 桥页（Doorway Pages, Bridge Pages）

又称为门页。作弊者针对不同关键词，制作大量低质量，甚至没有文字意义、只是充满关键词的页面，寄希望于这些页面获得排名带来流量。

很多时候桥页是由软件生成的，页面上的内容是没有意义的文字排列，用户根本没办法阅读。也正因为这些页面对用户毫无帮助，就不太可能在热门关键词中获得排名，只能是针对长尾关键词，所以需要大量甚至巨量的桥页才有意义。

桥页完全是以关键词排名和流量为目标，根本不考虑用户体验。

用户来到桥页后，又可以有两种处理方式。第一种方法是在页面最重要位置，比如顶部，以大字号链接到网站首页或其他网站（也就是作弊者真正要推广的网站）。用户在页面上看不到能读懂的内容，有很大一部分也就不得不点击这个链接来到其他网站，给作弊者带去有价值的流量。

第二种方法是页面自动跳转到真正要推广的网站上，比如使用 Meta 刷新或 JS 脚本跳转。通常这种跳转设置的时间是零，也就是用户来到页面上并没有看到桥页上的任何文字，就马上被自动转向到其他网站上。

8.2.10 跳转

前面提到过，搜索引擎接受度最高的是 301 跳转。在页面 URL 改变时使用 301 转向，搜索引擎会自动删除原来的 URL，把权重转移到新的 URL 上。

其他形式的跳转或转向，对搜索引擎来说就比较可疑。如 Meta 刷新，JS 跳转，使用 Flash 实现跳转等，其原因就是上面所介绍的桥页经常会使用跳转，所以连累到其他使用跳转的页面也会被搜索引擎怀疑是否有作弊的目的。

所以网站应该尽量避免使用除 301 之外的跳转。万不得已必须要使用时，跳转时间应该设置得长一点，而不能设置为零。比如 Meta 刷新，可以设置成这样：

```
<meta http-equiv="refresh" content="10;url=http://www.domain.com/">
```

用户打开页面 10 秒以后才跳转。这种有延时的跳转还有其他用处，比如用户注册网站账号或订阅电子杂志后显示的感谢页面就可以设置为 10～20 秒之后跳转到网站首页或其他合适的页面。

8.2.11 诱饵替换（Bait and Switch）

诱饵替换指的是作弊者先针对一些普通关键词制作页面获得排名后，页面再换成其他内容。一般有两种情况，一是先针对比较容易的长尾关键词制作页面，获得排名和点击后，把页面全部换成与热门关键词相关的商业价值更高的内容。第二种是先针对普通正当的关键词制作内容，获得排名后页面换成非法、成人、赌博等内容。

我们经常在搜索引擎结果中看到这种现象，页面内容完全改变后（很多时候这种改变并不是诱饵替换，而是正当的业务更改），原来的页面排名并不会立即消失。甚至在搜索引擎重新抓取页面新内容后，也不会立即消失，而是维持一段时间。由于搜索引擎有这种记忆特性，再加上诱饵替换页面被重新抓取就需要一段时间，所以诱饵替换页面往往能在原来正当或比较容易的关键词搜索中保持一段时间的排名。

8.2.12 关键词堆积（Keyword Stuffing）

正如这个名词本身所提示的，关键词堆积指的是在页面上本来没有必要出现关键词的地方刻意重复或者说堆积关键词，寄希望于提高页面对关键词的相关度或关键词密度，进而提高排名。

关键词堆积的地方既可能是可见文字，也可能是用户看不见的文字，如：

- 网页标题标签
- 说明标签
- 关键词标签
- 页面可见正文
- 图片 ALT 文字
- 页面内部链接锚文字中，尤其容易出现在页脚部分
- 页面 HTML 代码中的评论部分
- 隐藏在表格中等

关键词堆积在程度上有很大区别，所以也是一个比较模糊的作弊概念。有的网站属于轻度堆积，比如标题本来可以写成通顺自然的“童装批发零售”，却偏要写成“童装，童装批发，童装零售”。

有的作弊者则走向极端，在页面标题、说明标签、甚至页脚处的可见文字中列出几百个关键词。更有甚者，堆积的关键词与本页面内容还可能无关，本来是卖童装的页面，却加上更热门、搜索次数更多的 MP3、刘德华、小游戏等不相关关键词。作弊者的想法是搜索这些词的人数量巨大，希望自己撞大运能撞上几个。其实与页面内容无关的这种热门词，靠撞大运是撞不到的，反倒可能被搜索引擎认为是在作弊，连原本的内容都排不上去。

8.2.13 大规模站群

站群是很多 SEO 愿意使用的手法。自己制作一定数量的网站，既可以用来与其他网站交换链接，也可以用这些网站共同推一个主网站。由于站群是控制在自己手中，建立外部链接就容易多了。

站群的使用牵扯到一个度的问题，少量的网站，而且每个网站都有实质内容，搜索引擎不会因此惩罚或封杀。但是网站数量太大，因而网站质量必然很低时，搜索引擎就不能接受了。因为这样的站群唯一的目的就是影响搜索引擎排名，对用户已经没有什么价值了。大致上来说，几十个网站问题不大，但几百个甚至几千个网站，就会被认为是有作弊目的的站群了。

站群的准确判断有一定难度。SEO 人员为了避免被判断为站群，应该使用不同公司的主机，不同的域名注册信息，不同的网站模板，不同的网站内容，网站之间尽量减少交叉链接。

8.2.14 利用高权重网站

我们都知道域名权重对搜索引擎排名作用很大，有时候甚至大到了夸张的程度。英文维基百科在 Google 的疯狂排名就是这种机制的产物之一，不管用户搜索什么词，几乎总能在前面看到维基百科的条目。

很多黑帽灰帽自然会利用搜索引擎的这种算法特性为自己服务，自己的域名排名实在有困难时，利用别人的已经具备高权重的域名为自己服务。方式既可以是建链接到自己的网站，也可以直接用高权重域名的网页排名。

几年前，在维基百科建页面、建链接是不少黑帽乐此不疲的事。维基百科权重高，可以自由创建条目页面，自由加链接（链接长久保持在维基百科页面上难度很大，高级编辑和其他人都可能把链接删除），自由编辑，而且那时候导出外部链接没有 NoFollow。现在维基百科给所有导出到外部的链接加了 NoFollow，所以大部分作弊者没这么热心了。

被利用的高权重域名还包括那些免费但主流的博客托管商、图片分享网站、社会化网络等。很多人建立大量免费博客或社会化媒体网站账号，就是为了建设链接。

还有一种常用的黑帽手法是在大学网站上甚至政府机构网站上建网页。我们知道.edu 和.gov 域名往往有不一般的权重，黑帽 SEO 通过各种方式在教育或政府机构网站上建立一些网页，获得排名后再连到自己的网站，或者直接做转向把流量导入自己的网站。

这些作弊页面和大学或政府网站本身基本上没有任何关系。我们经常看到伟哥、类胆固醇药物、赌博、减肥、治脱发等高竞争度的关键词，排在最前面的是一堆大学的二级域名页面，这种页面 99%是作弊者放上去的，靠域名权重获得好的排名。当然，搜索引擎也一直修正算法，清理这些寄生在大学域名的页面。

怎样在大学域名上做网页呢？可以有很多方法。比如：

- 很多大学学生都可以在学校网站上自由创建个人网页，金钱、个人关系等就都可以派上用场了。
- 几乎所有大学老师都有在本校域名上的个人网页，很多教授的网页权重是相当高的。
- 有的大学网站可以自由发布消息，如分类广告、招聘信息等。
- 有的大学网站用的是现成的 CMS 系统， 这些系统有的允许未经认证的用户自由创建账户，有的有安全漏洞，如果你熟悉它的 CMS 系统，能找到方法创建账户。
- 当然还有的是直接黑进去了。

另外，有些大公司网站也并不像我们外人想象得那样纯洁无瑕。比如福布斯这样的大公司网站上就有莫名其妙的孤立的网页在宣传某种药物。怎么得来的呢？只有他们内部的人知道。

有时候利用高权重域名不一定是为给自己排名，也可能是为了打击对手，间接提高自己流量。举一个灰帽的例子。比如维基百科权重这么高，如果有些词你实在排不上去，而对手网站排在前面，就有人创建维基百科页面，然后在维基其他页面连向创建的新网页。维基内部链接并没有 NoFollow，链接权重传递通畅，内部链接是维基百科本身排名好的原因之一。

绝大部分站长自己的域名权重不高，但维基百科的页面做得适当的话，可以战胜其他绝大部分网站。维基百科的页面排名上升要简单得多，既打击了对手，也可以再想办法引导些流量到自己网站。

这个手法的坏处是，一旦你把维基页面做上去，自己想再超过它，也是难上加难了。

8.3 搜索引擎惩罚

从上面介绍的常见 SEO 作弊手法可以看出，黑帽与白帽之间的界限有时并不明显。看似作弊的手法，也很可能是因为疏忽或站长并不了解什么方法被搜索引擎认为是作弊。有的时候也可能是竞争对手有意陷害，比如垃圾链接，搜索引擎并不能百分之百正确判断出留在论坛、博客上的垃圾留言是谁制造的。再比如关键词堆积，多大数量、多高频率的堆积才被认为是作弊，也很难有个明确的标准。

搜索引擎也明白这一点，所以网站上出现单一涉嫌作弊的技巧，并不一定就会导致惩罚。不同作弊技术风险高低也不相同，导致的惩罚力度也不一样。

8.3.1 作弊的积分制

搜索引擎的作弊惩罚机制类似一个积分系统，每出现一个涉嫌作弊的地方，就给网站加一些作弊积分。当网站的作弊积分达到一定程度时，才给予不同程度的惩罚。

采取积分制另一个原因是，作弊的网站常常使用多种作弊手法，而不会只用一种。网站上出现一定程度的关键词堆积，或者偶尔链接向一个坏邻居，还都可以解释为疏忽。但如果又有关键词堆积，又有大量垃圾链接，还有买卖链接，网站还是站群的一部分，针对关键词又是那些热门又敏感的，就很难用疏忽和不了解来解释了。

举一个真实被惩罚网站的例子。

这是个卖家具的英国网站，站长发现他的网站被惩罚，Google 排名下降 40～60 位，所以在 Google 帮助论坛发帖寻求帮助。

Matt Cutts 很少在 Google 帮助论坛里回复，不过这次 Matt Cutts 发了一个回复，问发帖的站长是不是使用了不自然的建立外部链接的方法，如果有的话，应该特别注意。Matt 还特意举了个例子，是不是有赞助 WordPress 模板？这些赞助模板带来的链接被检测到以后，PR 和投票权重会被去掉，所以可能造成网站排名下降。

这个回复大概会让很多人心惊肉跳。如果赞助 WordPress 模板里面的版权链接就能导致网站被惩罚，那么就意味着可以给竞争对手赞助一堆链接，消灭竞争对手。当然，赞助模板和自己设计模板提供给他人使用还有着微妙的区别。赞助意味着付了费，变成付费链接。但如果这就是被惩罚的原因，Google 怎么知道模板是赞助的还是自己设计的？不可能知道。

Matt Cutts 的回复，促使我一直注意看这个帖子。后来的发展说明，设计或赞助 WordPress 模板本身并不是问题。

这个英国家具网站所赞助的模板很多是被用在了色情网站上。Matt Cutts 给出了很具体的例子，如图 8-1 所示。

图 8-1　WP 模板大量用于色情网站

随着讨论，这个网站被发现有个论坛，论坛里面有很多垃圾内容。站主的辩解是，论坛并不完全受自己控制。很多人在签名及会员资料页留下色情网站链接，他也没办法。我以前的观察是，论坛如果不注意删除这些明显有问题的内容，会伤害论坛本身。毕竟你不能说对论坛完全没有控制。不过这也不是主要原因。

另一位 Google 员工 JohnMu 后来又发了个回复，提醒那位站长，除了清理自己的论坛，也应该清理一下这位站长留在其他论坛、博客和新闻网站的垃圾评论。这些垃圾评论都是些“我同意”、“我不同意”、“顶”之类的话，留言人的名字是以关键词为锚文字的链接。典型的垃圾评论链接。

看来 Matt Cutts 和 JohnMu 说话都比较含蓄，点到即止。JohnMu 的评论让我意识到，实际情况恐怕和这位站长自称的完全不是一回事。检查一下这个站的反向链接，发现类似的垃圾博客、论坛评论数量很大，如图 8-2 所示。

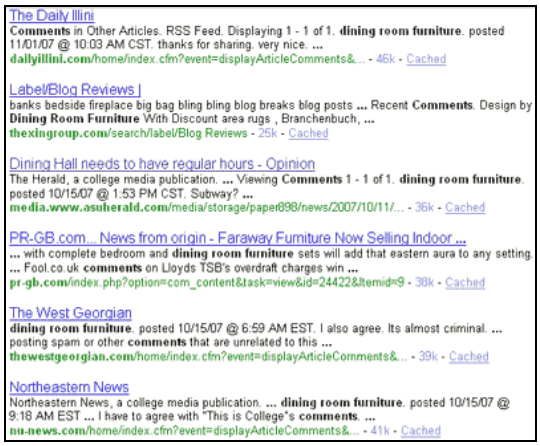


图 8-2　大量垃圾评论

同时我发现几个付费博客帖子，当然帖子里包含链接，比如图 8-3 所示的帖子。

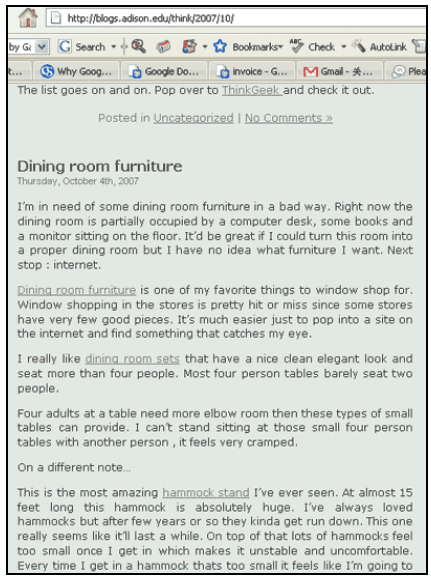


图 8-3　付费博客帖子及链接

怀疑其是付费帖子的原因是，这个博客一会儿谈助听器，一会儿说迪斯尼乐园，一会儿说 USB 硬盘，一会儿又推荐英国家具网站，完全没主题，谁也不挨谁。而且没有什么评论，显然也没有什么人看这个博客。

这些作弊手法单独出现，一般也不会有严重后果。适当做做页面优化，留几个博客评论什么的，无伤大雅。但如果把作弊的十八般武艺全用上，作弊积分达到惩罚门槛，就很难辩解说是无辜的了。

8.3.2　不要学大网站

这种积分系统的惩罚门槛也很可能不是固定的，而是一个可以滑动的范围，不同的网站有不同的惩罚门槛。SEO 人员必须理解的一点是，权重高的网站和成名网站能做的事，小网站不一定能做。经常在论坛中看到有人说，某某网站关键词堆积，也没见被惩罚，所以自己也要尝试一下。这对中小网站是很危险的。搜索引擎不是看不到大站的一些垃圾手法，只不过对大站的容忍度更高。

2006 年，德国宝马网站因为隐藏文字被 Google 删除。但这只是杀一儆百，告诉 SEO 行业搜索引擎对作弊的态度。宝马网站被删除后很快就做出了反应，修改网站，拿下隐藏文字，又很快被重新收录。很多人猜测，宝马与 Google 之间其实有我们所不知道的沟通渠道，所以才能这么快解决问题。对一个普通站长来说，网站被删除或惩罚，几乎没办法了解具体原因，不具备任何沟通渠道，所以还是不试为妙。

同样在 2006 年，纽约日报网站被发现使用了隐藏页面，但是也没有被惩罚。在这个意义上说，世界确实是不公平的。千万不能因为大站这么做，普通站长就跟着学习。

从某种意义上说，对大站名站容忍度更高是为了搜索引擎本身的用户体验，因为用户本来就期望看到大站的内容。如果搜索新闻内容时很难看到新浪、搜狐的页面，用户不会认为新浪、搜狐出了问题，而会认为这个搜索引擎不好，搜不到要找的新闻。

8.3.3 不要存侥幸心理

对搜索引擎来说，目标并不是百分之百消除垃圾和作弊，而是通过算法自动检测出大部分垃圾就足够了。只要搜索引擎用户不会在搜索结果中面对大量没有价值和意义的网站，就已经足够好了。所以一些网站采用作弊手法没有被发现，或者发现了也没有被惩罚。

但今天不惩罚，不意味着今后也一定不会被惩罚。搜索引擎反垃圾团队是质量控制很重要的一部分，他们的检测算法在不断改进中，一旦找到不会伤及无辜的检测和惩罚方式，搜索引擎对作弊就会毫不留情的下手。因为作弊和垃圾网站伤害的是搜索质量，这是搜索引擎黏住用户、保持和增加市场份额的根本所在，SEO 人员切不可有侥幸心理。

8.3.4 搜索引擎惩罚的种类

搜索引擎惩罚的形式很多，其中最容易判断的是整站删除。使用 site：指令搜索域名，如果网站完全没有被收录，就可以肯定是这几种情况：

- robots.txt 文件有问题，禁止搜索引擎收录。
- 服务器问题，使网站无法被搜索引擎抓取。
- 严重作弊行为被删除。
- 违法内容（如侵犯版权）被投诉后删除。

有的网站只是在搜索最主要关键词时被惩罚，其他次要关键词和长尾词排名不变。这种情况往往是外部链接优化过度或垃圾链接造成的，其中高度集中的锚文字是主要原因之一。

有的网站是所有的关键词全面排名下降。这里所说的下降是指大幅下降，比如从前十降到几十页以后。如果只是从第一页降到第二页，一般不是被惩罚，而很可能是算法变动或竞争对手优化得当排到了前面。

Google 也有的时候惩罚并不针对排名，而只是针对网站 PR 值。很多因为链接买卖而被 Google 惩罚的网站，都只是工具条 PR 值下降很多，但网站排名和搜索流量都没有什么变化。尤其出售链接的网站经常是这种现象。Google 的目的只是让他们没办法再靠 PR 值出售链接。

还有一种惩罚是排名下降固定数值。SEO 行业人员对一些名称一定都不陌生。比如负 6 惩罚，负 30，负 950 等，指的都是 Google 对一批网站排名下降固定名次。百度近一两年也有著名的 11 位现象。这里只简单介绍 Google 负 30 惩罚。

负 30 惩罚名称源自 2006 年 10 月一个站长在站长世界（webmasterworld.com）发的帖子。楼主发现他的一个网站很长时间排名第一，不过近几天排名降到 31，整整下降 30 位，并且排名稳稳的就停在那里了。不少跟帖的人也发现类似问题，都是原来排名第一的网页，下降整整 30 位。

有的人认为，这种负 30 惩罚和链接锚文字过度优化有关，有的人认为和 Google 的人工审查有关。

被称为小 Matt Cutts 的 Adam 在 Google 帮助论坛回答这个问题时说：

- 你是不是确认你的网站提供独特的内容？
- 大部分用户是不是觉得你的网站比其他网站更有用？
- 你的网站是否遵守了 Google 的站长指南？

有人继续问，Adam 说的这三条好像并没回答问题。Adam 再次强调说，这三条已经清楚地解释了这个问题。Adam 并没有承认也没有否认这种惩罚是否真实存在，而是直接列出了可能的原因，一般认为这是暗示负 30 惩罚是存在的。他列出的三条原因都是老生常谈，却很符合逻辑。

8.3.5 搜索引擎惩罚的检测

首先要明确的是，搜索引擎惩罚并不容易检测。网站一些关键词排名下降，流量下降，到底是因为被惩罚还是因为搜索引擎算法变动？或者有新的竞争对手加入进来？或者现有的竞争对手加强了 SEO？还是因为外部链接权重降低？这些情况之间很难准确区分。

下面几种方法可以参考，帮助站长进行判断。

1．使用 site：指令搜索网站域名

site:dunsh.org

如果没有任何结果，这是最简单的，网站被严重惩罚删除了。

2．搜索网站名称

在搜索引擎搜索网站名称，如果排在第一的不是你的官方网站，通常说明网站被惩罚了。当然这里所说的网站名称，不能是那种很宽泛的随意起的名称，像“计算机网”、“冶金网”之类的，而必须是真正的独特的网站名称或公司名称，如“点石互动”、“SEO 每天一贴”。

3．网管工具

Google、雅虎、微软必应都提供网站管理员工具，百度在 2010 年 4 月也开始内测网管工具。强烈建议所有站长都注册网管工具账号，验证自己的网站。这些网管工具会提供非常有用的信息，帮助站长判断网站是否有问题。尤其是 Google 网管工具，还会通知站长网站是否有作弊、病毒等情况。

4．搜索网站上特有的文字

和搜索网站名称类似，搜索一段只有自己网站上才有的特定文字，比如电话号码、邮件地址、备案号、联系地址等，如果排在第一的不是你的网站，说明网站被某种形式的惩罚了。

5．全面跟踪关键词排名

如果网站全部或大部分本来有排名的关键词全面大幅下降，很可能说明网站被惩罚。这里要强调的是“全面”记录跟踪关键词排名。很多时候网站一部分关键词排名下降，另一部分关键词排名上升或不变，这种情况一般并不是被惩罚。尤其是大中型网站，不同的关键词升升降降，一部分今天消失，另一部分明天上涨都很正常。只有在所有或大部分关键词全面排名下降时，才可能是惩罚。

6．检查日志

查看搜索引擎蜘蛛来访的次数、频率是否有变化。如果搜索引擎蜘蛛来访次数大幅下降，而网站本身规模和更新速度都没有变化，说明搜索引擎不再喜欢这个网站，很可能是因为某种形式的惩罚。

8.4 被惩罚了怎么办

经常在论坛里看到有人问，某某网站被惩罚了，大家帮忙看看是为什么。我也经常收到类似的邮件。

惩罚检测不容易，惩罚的恢复就更是一件头疼的事情。要想知道为什么被惩罚，纠正错误，恢复原有排名，必须非常清楚地知道这个网站以前做了什么？排名怎么样？流量怎么样？过去一段时间更改了什么东西？做了什么推广？惩罚的形式是哪种？这些详细情况都很难用一两段文字说清楚，所以给其他网站诊断惩罚问题是比较困难的。最了解自己网站的是站长自己。

下面这些方法可以参考。

首先，如果被惩罚的网站确实就是一个垃圾网站，那么原因很明确。只有做一个好

的网站，对用户有益的网站，才能解决根本问题。一个内容全是转载、抄袭，外部链接全部是博客论坛垃圾链接的网站，很难有好的排名，不值得费时费力。

如果你认为自己的网站不是垃圾网站，确实是对用户有帮助的网站，却被惩罚，可以尝试以下方法。

1．检查 robots.txt 文件

这是一个看似不可能，却常常发生的导致惩罚的原因。尤其是网站被全部删除时，更是要仔细检查 robots.txt 文件。不仅要人工查看代码，还要用 Google 网管工具验证是否有错误，造成禁止搜索引擎收录某些目录和页面。

2．检查服务器上其他网站

虽然搜索引擎一般不会因为站长使用的服务器上有其他作弊网站而惩罚服务器上其他网站，但是现在垃圾和作弊网站数目巨大，如果刚好有一个拥有大批垃圾网站的站长和你使用同一架服务器，服务器上大部分网站都是作弊和被惩罚的，那么你的网站也可能被连累。使用这个工具查看服务器上还有其他哪些网站：

http://whois.webhosting.info/

随机挑选一部分，查看一下收录和主要关键词排名情况。如果很多网站都有问题，那么尽早换主机提供商。

3．检查网站是否使用了转向

除了 301 转向，其他 Meta 更新，JS 转向都有可能被怀疑为作弊，哪怕站长的本意和想达到的目的其实与 SEO 作弊无关。如果网站上存在大量转向，尽快删除。实际上，一个设计得当的网站根本没有多少必要使用转向。

4．检查页面 Meta 部分代码

检查页面 Meta 部分是否有 noindex：

```
<meta name="robots" content="noindex, nofollow" />
```

和 robots 文件一样，这也是看似不会犯，但真的有人犯的错误。可能是公司其他部门人员加上去的，也可能是竞争对手黑进网站加上去的，还有可能是网站测试时加上去，正式开通却忘了删除。

5．彻底检查网站是否优化过度

查看页面是否有任何关键词堆积的嫌疑？是否为了加内部链接而加内部链接？是否锚文字过度集中？是否页脚出现对用户毫无意义、只为搜索引擎准备的链接和文字？有过度优化的地方，下决心“去优化”（减少优化）。很多关键词全面下降就是优化过度造成的，掌握优化的度是合格 SEO 人员必须亲身体验一遍的必经之路。

6．少安毋躁，切莫轻举妄动

如果确信自己网站没有作弊，遇到排名下降，切忌轻举妄动，不要忙着修改网站，先观察几天甚至几星期再说。排名下降不一定就是自己网站的问题，搜索引擎不断改变算法，有时推出新算法，监控数据表明新算法效果不好，过几天又改回去了。

有时排名下降正是搜索引擎对网站的考验，坚持按兵不动一段时间，搜索引擎就知道是个正常网站。一遇到排名波动网站就修改，反倒会引起搜索引擎的特别注意，这是一个刻意在做优化的网站。真正为用户而做，而不是为搜索引擎而做的网站，基本上无须关心排名波动，这种网站才会得到搜索引擎的青睐。

7．检查服务器头信息

虽然用户访问网站时看不出问题，页面正常显示，但搜索引擎访问时，服务器返回的头信息却可能有问题。我以前就遇到过客户的网站，访问完全看不出问题，但用服务器头信息检查工具查看时，返回的全是 404 代码，或者完全没有反应。

本书 SEO 工具部分介绍了检查服务器头信息的线上工具。Google 网管工具也刚刚增加了新功能，站长可以看到 Google 蜘蛛访问自己网站某个页面时抓取的内容，不仅可以查看头信息是否正确，还可以检查页面是否被黑客加上了病毒代码。

8．检查删除可疑链接

所谓可疑的链接包括：

- 大量交换友情链接
- 页脚上出现的只为搜索引擎准备的内部链接
- 买卖链接
- 连向坏邻居的链接
- 自己网站的大量交叉链接
- 与网站主题内容无关的导出链接等

这些链接看似不算严重作弊，但是与其他有作弊嫌疑的手法加起来，就可能使网站作弊分值达到被惩罚门槛。

9．检查是否有重复内容

包括网站本身不同 URL 上的相同内容，也包括与其他网站相同的内容。既可能是转载、抄袭造成的重复内容，也可能是技术原因造成的重复内容。如果一个网站从一开始就以转载、抄袭为主，被惩罚是应该的。加强原创内容是一个网站刚开始时必不可少的过程。

10．知道什么时候该放弃

被搜索引擎严重惩罚的网站，能恢复的可能性不高。不要说站长经常弄不清楚为什么被惩罚，所以也无法完全改正，就算很清楚为什么被惩罚，清除了所有作弊内容，也未必得到搜索引擎的原谅。有时候网站排名恢复是惩罚两三年之后的事了，有的时候根本没机会恢复。所以，很多时候尝试恢复还不如放弃，重新开始一个新网站。

第 9 章
SEO 专题

本章讨论几个经常遇到的 SEO 问题。

9.1 整合搜索优化

整合搜索是 2007 年年底 Google 首先推出的，现在已经被所有主流搜索引擎采用。

9.1.1 什么是整合搜索

所谓整合搜索（Blended Search），又称为通用搜索（Universal Search），就是在正常搜索结果页面中同时显示普通文字页面之外的图片、视频、新闻、博客、地图、图书等垂直搜索结果，如图 9-1 所示。

其实所有主流搜索引擎早就已经推出了垂直搜索，用户只要单击搜索结果页面上方的垂直搜索导航，就可以显示相应的结果。

图 9-1　Google 的垂直搜索选项

不过人都是懒惰的，搜索引擎大概发现绝大部分用户很少去点击垂直搜索导航，所以 Google 率先把垂直搜索结果整合入正常网页搜索结果页面，用户不再需要点击垂直搜索导航就能看到垂直搜索内容。

如图 9-2 所示就是一个出现三种整合内容的搜索结果页面。

排在最前面的是新闻资讯内容，中间有图片内容，最下面还有博客内容。

对搜索引擎来说，显示整合搜索结果需要解决几个难题。首先，搜索结果页面格式的安排。以前都是文字网页的时候，搜索结果格式整齐划一，用户也都熟悉了。现在要把视频、图片、新闻、博客等内容整合在同一个页面里，又不能显得混乱，还要实时处理，在网页的格式安排上要下一番工夫。

其次，不同类型的内容之间怎样比较相关性和权威度？以前返回内容都是网页，只要比较两个网页之间的相关性。但现在有了新闻、地图、博客、视频、图片等，视频和文字内容的相关性和权威性应该怎么比较？

图 9-2　整合搜索结果

再次，哪些搜索词会触发哪种整合结果？不太可能所有关键词都从所有垂直搜索中找结果。这需要根据具体的关键词，以大量用户数据为基础来判断用户的搜索意图。比如我们搜索“刘德华”时，图片也许还有视频内容，是必须要返回而且要多返回的内容，而地图内容就不重要了。

9.1.2　机会和挑战

整合搜索对 SEO 来说既是机会也是挑战。说它是机会原因在于，一些中小型网站在文字页面没机会排到前两页时，却有机会通过整合搜索结果进入前几名。进入资讯内容比较困难，通常被收为新闻源的都是大网站。除此之外，视频、图片、博客、地图等内容难度就比较小，大网站很可能还没有时间、精力顾及到这些垂直领域内容，中小型网

站多花一些时间和精力，做得更专门化，就有机会在比较热门的搜索结果中出现。

统计表明，虽然整合内容出现次数还是比网页少得多，但竞争也小得多，图片、博客、视频、地图等内容出现在第一页的概率比网页高出 10 倍以上。据 2009 年 Forrest 的研究（http://blogs.forrester.com/interactive_marketing/2009/01/the-easiest-way.html ），超过一半的 Google 搜索结果中有视频内容，视频结果出现在搜索结果页面第一页的概率比普通文字页面高 50 倍。

说它是挑战是因为，传统 SEO 要优化的只是普通页面，以文字为主。现在 SEO 的工作范围更加广泛，各种多媒体内容和社会化媒体内容都要被纳入 SEO 的范围。SEO 人员必须从战略思考上提高一个层次，不再局限于传统意义上的页面优化。

自推出整合搜索以来，Google 搜索页面出现整合结果的数量越来越大。相对来说，百度整合内容数目还很少，除了新闻内容经常出现以外，其他很少看到，在搜索带有地名的关键词时，有时会出现地图内容。

下面简单讨论几种常见垂直领域的优化技巧。

9.1.3 新闻搜索

进入新闻（资讯）搜索最重要的是被搜索引擎纳入新闻源。要成为搜索引擎新闻源，对中小型网站来说还比较困难。有条件感兴趣的站长可以参考提交新闻源页面。

Google：http://www.google.cn/support/news_pub/bin/request.py?contact_type=suggest_content

百度：http://news.baidu.com/newsop.html

被搜索引擎纳入新闻源，对整个网站所有页面的普通排名都有好处，能够提高网站权重，提高页面爬行频率。

搜索引擎返回资讯内容时，先选定哪几条新闻应该被显示，如图 9-3 中的三条新闻，这主要靠用户关注度及媒体报道数量判断。再选择应该显示哪个网站的新闻内容，如图 9-3 所示，每条新闻都有几个到几千个来自不同网站的报道，搜索结果中只能显示一个，用户点击“xxx 篇相关文章”才能看到不同网站对同一条新闻的报道。

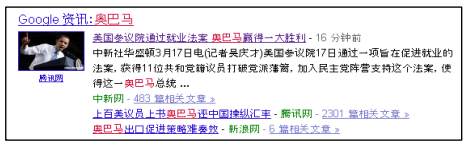


图 9-3　Google 的新闻整合结果

新闻内容的排名与普通网页的排名算法不同，需要注意以下几点：

- 发布和更新时间。由于新闻内容本身的特性，越新发布的内容越有排名优势。

- 原创。转载其他新闻源内容，效果没有原创那么好。就算转载内容的发布时间更新，如果没有附加价值，排名也没有发布时间较早的原创高。
- 地域相关。很多新闻有地域性，同一个地区的新闻源网站具有优势。比如说上海地区发生的新闻，上海的新闻源网站就更有排名优势。
- 标明时间、地点。新闻稿中最前面标明时间、地点是标准的新闻写作格式，如“中新社华盛顿 3 月 17 日电”，也有助于搜索引擎判断新闻内容的地域性和实效性。
- 网站权重。新闻源网站本身的权重也是重要因素。这里所说的权重与 PR 值或其他来源于外部链接的权重不同，更多的指的是与主题相关的权重。比如说大量发布金融新闻内容的网站，在金融领域就会有更高的权重。
- 点击率。如果说普通页面点击率影响排名只是猜测或影响很小，新闻内容的点击率则有明确的直接影响。同样的新闻内容，点击率高的那一条排名会更快上升。
- 图片。新闻页面中的图片经常会被显示在搜索结果中，有助于吸引目光，提高点击率。
- 其他页面因素。与普通页面相似，页面标题、URL 及正文出现目标关键词，也都会计入新闻排名当中。

9.1.4 图片搜索

图片内容是现在搜索结果中最常出现的内容之一。

为了让自己网站的图片内容出现在整合搜索中，应该注意下面几点：

- ALT 文字。这是图片优化最重要的部分，因为 ALT 文字本身就是为了说明图片内容。ALT 文字应该出现目标关键词，准确描述图片内容。
- 页面标题。图片所在页面的标题也很重要。标题写法参考前面页面优化部分。
- 图片周围文字。图片出现的周围文字内容也在一定程度上表明图片内容。当然，这里所说的图片周围文字指的是 HTML 代码中图片文件的前后，而不一定是用户所看到的页面上图片周围的文字。有时候由于排版原因，视觉上看来处于图片周围的文字，在代码中可能相距很远。
- 图片质量。图片的解像度和像素，也就是图片质量也影响图片优化。质量越高，被排到前面的可能性越大。搜索引擎并不想返回低质量的图片。
- 导入链接锚文字。导入到图片出现页面的链接所使用的锚文字，既表明页面本身的内容主题，对图片内容也是个说明，会被搜索引擎考虑在内。

9.1.5 视频搜索

如图 9-4 所示是 Google 搜索结果中出现的视频内容。

视频文件既可以放在自己网站上，也可以放在如 Youtube、酷 6、六间房、土豆等视频分享网站上。在实际使用中，大家会发现，出现的视频内容绝大部分来自视频分享网

站和一些门户级的视频内容，很少有企业或个人网站上的视频，原因是这些视频分享网站权重往往远高于普通网站。

图 9-4　Google 的视频整合结果

所以，通常优化视频内容需要把视频文件放在视频分享网站上，得到的搜索流量其实是对视频网站的访问，而不是自己的网站。这就需要 SEO 人员再想办法把用户从视频网站引到自己的网站上，比如在视频中加入片头、片尾，版权字幕，或者视频本身就包含品牌信息。

视频页面的优化需要注意以下几个方面：

- 标题。视频文件所在的页面标题，通常就是提交视频时自己撰写的视频标题。和普通页面一样，应该嵌入目标关键词。
- 文字说明。与标题类似，提交时填写的文字说明会显示在页面的视频说明部分，也需要包含相应关键词。由于篇幅允许稍长，可以加入关键词的变化形式。
- 播放次数。几乎所有视频网站都会显示本视频的播放次数，次数越多说明视频越受欢迎，搜索引擎给予的权重也就越高。
- 用户评分。同样，几乎所有视频网站都允许用户给视频评分，并且显示在视频页面上。分数越高，视频内容质量高的可能性也越大。
- 用户评论及留言。主要是看留言数目，播放次数有时还可以作假，留言数目作假难度就大大提高了。播放次数、评分及留言，都在一定程度上表明视频的受欢迎程度，有点类似于页面的导入链接。
- 链接。和普通页面一样，有越多外部链接，视频页面权重越高。与文字页面相比，视频排名需要的链接数就少多了（至少目前是这样）。
- 标签。提交视频时，系统通常要求用户填写几个说明视频内容的标签。在主题相关的情况下，尽量多填写几个相关标签，这样提交的视频出现在其他视频的相关视频里的机会就更多，被看到、被评分、留言的机会都更多。
- 视频嵌入。视频分享网站都提供代码，允许站长把视频嵌入到其他网站。鼓励其他网站嵌入视频，因为都会计入播放次数，提高能见度。
- 缩略图。一张好的缩略图不仅吸引视频分享网站上的用户，也吸引搜索引擎的用户。

9.1.6　地图搜索

图 9-5 所示是 Google 出现地图内容时的搜索结果页面。

图 9-5　Google 的地图整合结果

可以看到，地图结果在视觉上非常突出，占了页面很大篇幅，原来排第一位的文字网页都几乎都被推到第二屏，所以对某些关键词，出现在地图结果中显得非常重要。

- 注册验证。要出现在地图结果中，首先需要在 Google 本地商户中心注册，注册地址是：http://www.google.com/local/add?hl=zh-cn&gl=cn。然后通过手机或信件验证地址。完成验证后，提交的商户内容才会出现在地图结果中。
- 使用关键词。提交内容时，公司名称及说明文字中自然都要使用目标关键词。
- 上传图片和视频。提交本地商家内容时，注册人还可以上传图片及视频，展现公司形象。这些图片和视频与关键词的相关性倒在其次，更重要的是注册人花时间上传视频和图片，显示了认真和可信程度。
- 标签。登录 Google 账号，在“我的地图”部分可以给不同的地图地址加标签。SEO 人员可以自行在提交过的地址上标注上自己的企业名称。多一些人标注当然效果更好。
- 本地目录及黄页登录。把公司网站提交到黄页及带有本地性质的网站目录中，只要 Google 收录了这些网上黄页和目录页面，就会把公司与相应的地址联系起来，对提交的本地商户信息是个再次验证和加强。
- 评论。提交内容在 Google 地图中出现后，允许所有用户进行评论。评论越多，说明关注的人越多，也代表了某种受欢迎程度。
- 其他网站出现公司地址。除了上面说的黄页及本地目录外，在其他任何网站如分类广告、合作伙伴网站等地方出现公司名称、网址及联系地址时，都可以使 Google 更加确认所提交的公司与地址之间的联系真实性，使你提交的内容在地图排名中提前。

9.2　更改域名

一般来说，在网站策划阶段就应该选择一个最恰当的域名，一旦确定域名，就不要

轻易改动。但有时更改域名也可能是万不得已的。比如：

- 发现了更好的域名。策划网站时最中意的域名被注册了，后来发现人家没续费，自己又拿到这个域名，或者花钱把域名买过来。
- 公司合并或改名。这种商业决定的重要程度当然远高于域名的选择，不可能为了保留域名，公司就不进行合并。
- 法律问题。侵犯了其他公司/人的注册商标等。

一些著名网站有过更改域名经验，如美味书签 Delicious 网站从 del.icio.us 改到 delicious.com。国内著名的 IT 社区 5Gsns.com 推出不久又改域名为 5Gme.com。

由于更改域名必然使所有 URL 发生变化，从 SEO 角度来说需要做恰当的处理。

（1）全站做 301 转向，旧域名上的所有页面（不仅是首页）全部按原有目录及文件格式转向到新域名。如 URL：

http://www.domain.com/example/page.html

做 301 转向到：

http://www.newdomain.com/example/page.html

这样旧域名的权重会转移到新域名。

（2）尽量把指向旧域名的外部链接改到指向新域名。这是一个不易完成的工作，因为绝大部分外部链接是来自其他人的网站，不是自己所能控制的。可以通过流量统计和外部链接查询工具，找到能带来点击流量及比较重要的外部链接，联系对方站长说明情况，把链接指向新的域名。尽最大可能，能联系多少就联系多少。

301 转向虽然能比较好地解决 URL 变化问题，但并不能传递 100%的链接权重，每一个 301 转向都造成链接投票力的损失。而且搜索引擎识别 301 转向并重新计算权重需要比较长的时间，通常要几个月。依靠搜索引擎自己判断，不如自己把问题解决，尽量减少不可控因素。

（3）保留旧域名，并一直保留 301 转向，除非因为法律问题不能再持有旧域名。一部分指向旧域名的外部链接是永远不会改到新域名的。只要旧域名和 301 转向一直存在，这些链接还会传递一部分权重到新域名，不至于全部浪费。

最后要提醒的是，不到万不得已，不要更改域名。域名的历史和信任度是没办法完全传递的。

9.3 多个域名的处理

多个域名对应一个网站是常见现象。原因包括：

- 为保护公司品牌注册多个类型的域名，以及各个国家的地区域名。比如主域名是 abc.com，公司还可以注册 abc.net，abc.org，abc.co.uk，abc.sg，abc.com.cn 等。
- 类似和错拼域名。为了不使用户混淆，也减少可能发生的骗局、名誉损害等，很多公司还会注册或购买与公司主域名类似或错拼的域名。如 Google.com 是主域名，还可能注册 goolge.com，gooogle.com 等。
- 闲置域名。很多做域名买卖的人手中有大量闲置域名，一些闲置域名还有用户直接输入网址访问。与其让闲置域名无法解析与访问，不如利用起来指向有内容的网站。

上面几种情况都属于合理使用多个域名。但有一种说法是对多域名的误解，那就是有的 SEO 和站长认为，多个域名直接解析到同一个网站，也就是说多个域名都可以访问，返回的是一模一样的网站，搜索引擎会给予多个网站排名。这是不会发生的。

将多个域名不做任何处理地解析到一个网站，最好的情况是搜索引擎只返回其中一个网站给予排名。如果运气不好，可能因为复制内容和疑似作弊，所有域名都被惩罚。

处理多个域名对应一个网站方法很简单，选择其中一个最好的域名当做主域名正常解析，其他所有域名全部做 301 转向到主域名。

检查 goolge.com，gooogle.com 这类域名的服务器头信息就可以看到，都是 301 转向到 google.com：

```
#1 Server Response: http://www.gooogle.com
HTTP Status Code: HTTP/1.0 301 Moved Permanently
Location: http://www.google.com/
```

9.4 更换服务器

更换服务器也是网站经常遇到的问题。有不少站长询问更换服务器会不会影响网站排名。只要操作得当，网站转移到另一架服务器上，不会对排名有任何影响。除非你倒霉地把网站转移到了一个已经被搜索引擎惩罚的 IP 地址上。整架服务器 IP 被搜索引擎惩罚的概率非常低。

正确的服务器转移过程是：

- 做一个完整备份，包括数据库和所有程序、页面、图片文件，以防万一。
- 将 DNS（域名服务器）TTL 设置为很短，如几分钟。TTL 控制 DNS 服务器的缓存时间，设置为很短时间，所有 ISP 及搜索引擎将随时查看 DNS 信息，而不是使用缓存中的 IP 地址。
- 注册新主机账号，上传文件。
- 确认所有文件在新服务器上一切运转正常后，更改域名服务器，将网站解析到新服务器 IP 地址，同时旧服务器上的网站保持运行。一些用户贡献内容的网站，

可以考虑旧服务器上暂时关闭用户发布新内容，以防万一转移过程出现问题，可能会造成新发布的数据丢失。

- 在新服务器上检查日志文件，确认搜索引擎蜘蛛开始爬行新服务器上的页面后，说明转移已经完成。域名解析理论上最长需要三天时间，当然为保险起见，也可以再多等两三天，确保所有搜索引擎蜘蛛都知道网站已经转移到新的服务器和 IP 地址。
- 确认解析过程完成，旧服务器上已经没有任何用户及搜索引擎蜘蛛的访问后，旧服务器账号可以关闭，转移过程完成。

整个服务器转移过程中，网站一直是可以正常访问的，因而不会对爬行、收录、排名有任何影响。

9.5 用户行为影响排名

前面讨论过，SEO 的根本原理在于提高网站内容相关性、权威性和实用性。其中的实用性指的就是用户行为，也就是说网站实用性越高，对用户越有用，用户表现出来的行为又会影响网站排名。

9.5.1 用户行为信息收集

有一个很明显的现象，所有主流搜索引擎都提供浏览器工具条，用户不必到搜索引擎网站进行搜索，只要在工具条中输入关键词，就可以直接来到搜索结果页面。所有搜索引擎都花大钱大力，希望成为浏览器的默认工具条提供者，比如给浏览器开发商搜索广告回扣，或者与电脑厂商达成协议，在出售的电脑上预装工具条。浏览器和电脑提供商与哪家搜索引擎达成合作，换用哪家搜索引擎工具条，都是搜索行业的新闻。

搜索引擎不遗余力地推广自己的工具条，其目的不仅仅在于提高自身的搜索市场占有率，另一个重要意义是搜集、记录用户搜索及访问网站的行为。只要浏览器工具条启用，用户搜索了哪些关键词，访问了哪些网站，在网站上访问了哪些页面，停留多少时间，搜索引擎都可以通过工具条进行数据统计。

当然，用户访问时的行为方式是一种噪声比较高、作弊可能性也比较大的数据。搜索引擎在使用这种数据作为排名因素时都会非常小心。Google 工程师多次强调过，工具条数据基本上不会用于排名，但并没有排除这种可能性。

实际上大部分 SEO 专家认为，用户行为数据已经或多或少地被用在排名上，而且在未来几年用户行为所占的比重还会越来越高，SEO 人员必须重视。

9.5.2 影响排名的用户行为

可以影响排名的用户行为如下。

1．网站流量和 Alexa 排名

这是最直接、误差也最大的因素。尤其是 Alexa 排名，由于样本分布不均匀，又容易作弊，与网站真实流量往往相差很大。不过网站总体流量毕竟在一定程度上说明网站的受欢迎程度。

2．点击率

网站页面在搜索结果中被点击的比例。虽然点击率主要是由排名先后决定的，但是由于网站的知名度、用户体验、标题写作等方面的差别，有的网页也可能排名靠后，却有比较高的点击率。所有搜索引擎都记录搜索结果的点击率，发现某些页面有超出所在排名位置应该有的平均点击率，可能认为这些页面对用户更有用，因而给予更高排名。

这也就是所谓的“点击器”对网站排名有时有效果的原因，尤其是对百度。不过使用过点击器的很多人会发现，通过模拟点击提高排名效果并不持久，说明点击率很可能是排名因素之一，但搜索引擎也有防范措施。

3．网站黏度

弹出率、访问时间、访问页面数这些反映网站黏度的指标，也都可以被工具条记录。弹出率越低，停留时间越长，访问页面数越多，说明网站用户体验越好，可能对排名有正面影响。

4．用户特征

不管是新用户还是老用户，一个用户多次返回访问同一个网站，说明这个网站有用处，应该给予更好的排名。比如用户本身是哪个领域的专家，他也可能对所访问网站的排名有影响。

5．重复搜索以及相应的点击

用户第一次搜索，点击一个网站之后，可能没有找到有用的答案。单击“返回”按钮再次来到搜索引擎搜索结果页面，点击另一个网站，才找到对自己有帮助的结果。这种行为模式可能影响这两个网站的排名。

6．品牌名称搜索

如果某个品牌名称或域名本身被很多用户搜索，这个品牌的官方网站就很可能是用户在寻找的网站，排名可能相应提高。比如在 SEO 领域，“点石互动”具有品牌效应。虽然搜索“SEO”时点石网站排名不在第一，但是有很多人直接搜索“点石”或“点石互动”，也会对点石网站排名有影响。我们可以看到，“点石互动”首页标题中完全没有出现“SEO”这个词，但搜索“SEO”时排名一直不错。

7．社会媒体网站

社会化媒体网站出现页面的链接或网站名称，也可能影响网站排名。这一点下面再比较详细地探讨。

9.5.3 回归用户体验

我们可以想象这样一个场景，一个用户自身拥有 SEO 网站，订阅大量 SEO 博客，这一点搜索引擎可以通过流量统计服务和 RSS 阅读器判断。这个至少对 SEO 比较关注、很可能是 SEO 专家的用户，在搜索引擎搜索“SEO”点击第一个结果后 10 秒钟就按返回按钮重新来到搜索结果页面，点击了第三个结果，在这个网站上访问了 10 个页面，停留了 30 分钟，然后每过几天，这个用户都会再访问排在第三名的网站，并停留较长时间，访问多个页面。这样一种行为模式如果多次重复，对原本排在第三位的网站很可能有正面影响。

除了工具条外，主流搜索引擎还开发了很多其他服务，都能收集用户访问数据，比如百度、Google、雅虎等都提供网站流量统计服务、搜索广告联盟等。搜索引擎之所以免费提供这些服务，肯定不是一无所图，他们从中得到的数据既可以用来了解互联网整体趋势，也可以用来了解用户行为方式。

在这个意义上，要做好 SEO，还是要回归根本，那就是提供好的用户体验，提高用户忠诚度，给用户提供好的服务。这既是做网站的基本，也是 SEO 的未来。

9.6 域名与 SEO

域名是互联网公司以及个人站长最重要的无形资产之一。所有网站内容、流量都与特定域名相联系。一个好的域名对 SEO 及网站运营都有一定的影响。

9.6.1 域名后缀

SEO 行业曾经流行这样一种观点：.edu 和.gov 等不能随便注册的域名天生有更高的权重。.edu 只有美国大学等教育机构才能注册，.gov 只有政府部门才能注册。注册限制决定了这些域名很少能被用来做垃圾网站，因而搜索引擎给予排名优势。还有人认为.org 域名比.com 域名更有排名优势。

这种说法有一定道理，但并不准确。.edu 和.gov 的权重和排名优势并不是天生的，而是由于这种域名上的内容往往质量确实比较高。而.org 域名比.com 域名更有 SEO 优势，则没有什么确实根据。

还有人认为.info 域名权重比其他域名低，这也没什么根据。我们在搜索结果中很少

看到.info 域名，只是因为用.info 域名认真做站的公司和站长本来就很少。

SEO 行业公认.edu、.gov 网站是最好的外部链接资源，其原因也就在于这些网站通常质量比较高，权重比较高，但是这些权重并不是天生的。

所以域名后缀的选择并不是很重要，在可能的情况下还是以.com 为最好，因为这是大部分用户最习惯的域名，如果不提是哪种域名，用户都会认为是.com 域名。

9.6.2 域名年龄

域名注册越早，对排名越有利。正因为如此，购买老域名是 SEO 行业的常见做法。如果你有一个 20 世纪 90 年代就注册了的域名，这绝对是个宝，做绝大部分关键词相对新域名都有很大优势。注册一些不错的域名然后养起来，放上一些简单页面，过几年有了更充实的内容、更好的计划，再正式做网站，是一个不错的选择。

9.6.3 域名第一次被收录时间

除了域名注册日期以外，域名上的内容第一次被搜索引擎收录的时间也很重要。有的老域名注册以后没有解析，所以没有被搜索引擎收录任何内容，其年龄优势就比不上很早就被收录内容的域名。

虽然我们并不能查询到特定域名到底什么时候开始被搜索引擎收录，但可以使用前面竞争对手调查部分介绍的互联网档案馆查询网站历史内容。互联网档案馆中第一次出现域名内容与搜索引擎第一次收录的时间未必完全一样，但是有重要的参考作用。

9.6.4 域名续费时间

Google 在 2003 年 12 月申请的一份名为“基于历史数据的信息检索”的专利中提到，域名续费时间可以作为排名的因素之一。其逻辑在于，续费时间长，说明站长对网站认真，不太可能把域名用来做垃圾网站。通常做垃圾网站的黑帽 SEO 只会注册域名一年，看看效果如何，带不来流量或被惩罚，域名也就放弃了。

当然，专利中提到的内容是否真实地运用在排名算法中，谁也不知道。但真正认真的公司和站长，把域名多续费几年也没有任何损失。当年 Google 的这份专利刚被发现时，站长世界的创始人 Brett Tabke 马上把 webmasterworld.com 域名续费了十年，可谓对 SEO 的嗅觉非常敏感，也为其他 SEO 做了一个很好的提示。

9.6.5 域名包含关键词

这一点主要适用于英文网站。

域名中包含目标关键词，是目前搜索引擎排名的一个比较重要的因素。原因有两方面：一是域名中的关键词本身就给予了排名权重，二是很多网站转载文章时留的链接就是原文 URL，如果域名中包含关键词，那么相当于链接锚文字中就包含了关键词。前面提到过，锚文字是页面排名的重要因素之一。

9.6.6 连词符使用

四五年前，曾经很流行在域名中使用连词符，甚至多个连词符。前面提到，域名中包含关键词对 SEO 有些好处，但是这种好的域名一般早就被注册过了，很多 SEO 就转而注册用连词符把关键词分开的域名。比如 k1k2.com（k1 和 k2 是关键词）早就被注册了，站长转而注册 k1-k2-k3.com。

现在不建议注册这种使用连词符、包含关键词的域名。因为特殊原因包含一个连词符，问题不大，包含两三个，则可能弊大于利。第一，连词符给用户的印象不好，很容易让人联想到垃圾甚至骗局网站。我们很少看到大公司大网站使用带连词符的域名。第二，搜索引擎对域名中包含多个连词符也变得比较敏感。虽然不至于直接惩罚，但很可能使网站的可疑度又增加了一点。

9.6.7 品牌优先

在域名中包含关键词，其实是仅仅考虑了 SEO 因素，而没有站在更高的视野。我们可以看到一个规律，真正的大品牌都不含有行业关键词。比如英文搜索引擎领域最大的品牌是 Google，中文搜索引擎是百度。“Google”、“百度”这两个词都与“搜索引擎”毫无字面关系。最大的搜索引擎并不是 searchengine.com。最大的英文网上书店是 amazon.com，而不是 books.com。最大的中文网上书店是 dangdang.com，“dangdang”或“当当”与“书”也没有任何关系。

在某种意义上说，品牌名称与行业或产品名称相距越远反倒越好。

互联网公司还有一个明显的现象，大品牌很多是硬造出来的词，品牌名称诞生之前就不是一个词，更不要说是关键词了。Google，Youtube，Twitter，Facebook，等等。

而品牌名称一旦确定，域名也就随之确定。所以如果你想创造一个真正的大品牌，域名中反倒不要包含关键词。

9.6.8 域名长短

要想在域名中包含目标关键词，往往就得使用比较长的域名。在很多情况下，长域名所带来的 SEO 优势，远远比不过带来的品牌和用户体验方面的弊端。所以在可能的情况下，域名越短越好，不必为了包含关键词而使用一个超长的域名。

关键词与品牌相比较，永远是品牌优先，用户易用性优先。短域名易记、易写、易传播，由此带来的好处通常会超过域名中含有关键词的好处。当然，又包含关键词又短，那是再好不过了。

9.6.9 域名买卖历史

域名注册以后是否曾经转手？域名主人是否与垃圾网站有关联？这对域名权重也有一定的影响。所以在买二手域名时，应该注意查看一下域名曾经转手多少次，主人是谁，互联网图书馆中收录的内容是什么，有没有与敏感内容相关联。如色情，赌博，药品买卖等这些领域，经常作弊泛滥。

9.6.10 匿名注册信息

为了逃避搜索引擎检查域名注册人信息、转手历史，有的人使用匿名注册信息。这是一柄双刃剑。虽然搜索引擎无法从域名注册信息判断这是谁的网站，但是也很可能在可疑分数上增加了一分。一个正常认真的网站，最好还是使用真实注册信息。

9.6.11 域名权重

域名权重是一个很宽泛的概念，与域名年龄、内容原创度、网站规模、外部链接数量质量都有关系。权重高的域名往往做任何关键词都能很快看到效果，无往而不利，如英文维基百科。连很多普通用户都注意到，不管搜索什么关键词，几乎总能看到维基百科条目排在第一页。当然，要达到这样的权重，就不是单单 SEO 所能做到的了。

9.7 主机与 SEO

所有的网站都依附于服务器硬件。服务器（或者说主机）性能也会影响 SEO 效果。

9.7.1 IP 及整个服务器惩罚

一个 IP 地址或整个服务器被搜索引擎惩罚是很少见的情况，除非这个 IP 地址上的大部分网站都因为作弊被惩罚，这种情况下没有作弊的网站才可能受连累。不过使用虚拟主机时，同一架服务器刚好碰上大部分网站都作弊的可能性非常低。通常黑帽站长租用整台服务器，把自己的作弊网站都放在同一架服务器上，才会发生整个 IP 和服务器被连锅端的事情。

使用虚拟主机并不是必然比租用整台服务器效果差，网络上大部分网站都是使用虚拟主机的。搜索引擎并不歧视虚拟主机用户。

9.7.2 服务器设置

有的主机设置有问题，整个服务器都禁止搜索引擎爬行，普通用户访问网站时则没有问题。还有的服务器 404 错误信息设置不正确，页面不存在时，用户看到的是 404 错误信息，但返回的服务器头信息却是 200，这样会使搜索引擎认为存在很多复制内容。SEO 人员需要检查日志文件，确保搜索引擎蜘蛛能够顺利爬行，并且返回正确的头信息。

9.7.3 稳定性

服务器三天两头死机，必然会影响搜索引擎爬行和收录。轻者不能及时更新页面内容，抓取新页面，重者搜索引擎会认为网站已经关闭，或者爬行频率大大降低。好在主机稳定性如何是非常直观的，经常打不开网站的站长最好及时换其他主机服务商。

9.7.4 主机速度

除了影响网站用户体验和转化率，主机速度也严重影响网站收录。对一个给定网站来说，在搜索引擎眼中有一个固定的权重，搜索引擎会分配一个与权重匹配的相对固定的爬行总时间。网站权重越高，分配的总爬行时间越长。如果网站速度比较慢，搜索引擎抓取一个页面就需要更长时间，这样就会影响总收录页面数。

这对小网站还问题不大，对大中型网站来说，页面调用时间长，抓取占用时间就长，必然影响整个网站所能收录的页面数量。而对大中型网站来说，提高页面收录率是 SEO 的最重要工作之一。在这个意义上，主机速度直接影响 SEO 效果。

除了影响网站收录，页面打开速度也是影响 Google 排名的直接因素之一。2009 年 Google 反垃圾组带头人 Matt Cutts 特意提醒 SEO 人员，Google 很有可能在 2010 年把页面速度作为排名因素之一。2010 年 4 月，Google 公开宣布速度正式成为排名因素，而且已经试用了几个月了。普通用户和 SEO 人员并没有注意到这个变化，因为只影响很小一部分搜索结果，速度只是很小的一个因素，权重及相关性还是最重要的。

百度是否考虑页面速度并不明朗，只是声明需要非常谨慎，但显著影响到用户体验的广泛要素，最终都会被搜索引擎谨慎地尝试和使用。

9.7.5 URL 重写支持

现在的网站都使用的是 CMS 系统，数据库驱动，所以将动态 URL 重写为静态是 SEO 必不可少的工作。URL 重写需要主机支持。有一部分站长对虚拟主机有误解，认为虚拟主机不支持 URL 重写。其实虚拟主机可以完美支持 URL 重写。如果你使用的主机不支持，这只是主机提供商没有安装相应模块而已，并不是说虚拟主机就不能支持。

要支持 URL 重写，LAMP（Linux+Apache+MySQL+PHP）主机需要安装 mod_rewrite 模块，Windows 主机可以使用 ISAPI Rewrite 等模块。

9.8 多语种内容

随着外贸行业不断发展，发布多语种网站内容也越来越常见。很多国内企业把外文网站当做重要的营销渠道。不同语种的网站在 SEO 原则上是相同的，熟练掌握中文 SEO，也可以把同样的手法用在其他语种的网站上。

不同语种内容不会造成复制内容，哪怕同样的文章从中文翻译成英文、法文、日文，这些不同语种页面对搜索引擎来说也是不同的内容。

要注意的是，尽量不要在一个页面中混合不同语言。有的网站把英文内容紧接着放在中文内容下面，这样不能达到最好效果。应该把不同语种内容完全分开，中文页面就是中文页面，英文页面就是纯英文页面，不然搜索引擎在判断页面语种时会混淆。

9.8.1 多语种页面处理

不同语种的页面在网站上可以有三种处理方式。

（1）不同语言网站完全独立，放在不同的国家域名上。如中文版就放在 abc.com..cn，美式英文版放在 abc.com，日文版放在 abc.co.jp，英式英文版放在 abc.co.uk。这样做的好处是用户及搜索引擎都能轻易分辨语种和地理位置。不同语种的页面写作、内容安排上也都可以自由发挥。不同国家域名的网站可以放在相应国家的主机上，有助于关键词排名的地理定位。

缺点是这些网站完全独立，在推广上需要花更多精力，需要同时优化和推广多个网站。

（2）不同语种网站放在主域名的子域名上。如中文版本作为主网站，放在 abc.com，英文版本放在 en.abc.com，法文版放在 fr.abc.com。这样做的好处是子域名多少可以继承一点主域名的权重。如果主域名已经存在较长时间，有比较高的权重，对多个子域名的优化排名都有好处。用户和搜索引擎也可以轻易分辨出这些子域名是不同语言版本。不同子域名也可以放在相应国家的主机上。

缺点和使用不同国家域名一样，不同子域名在搜索引擎眼里基本上是不同的网站，也需要同时优化推广多个网站。

（3）不同语种网站放在主域名的二级目录下。如中文版主域名为 abc.com，英文版放在 abc.com/en/下面，法文版放在 abc.com/fr/下面。这样做的优点是二级目录完全继承主域名权重，在搜索引擎眼里，不同语言版本实际上是一个网站。如果你已经有一个权重相当高的主域名，在二级目录下推广其他语种页面就容易得多。

缺点是用户和搜索引擎都可能对网站语种产生一定混淆，而且不同二级目录很难放在不同国家的主机上，技术上实现起来比较困难，这对关键词地理定位很不利。

9.8.2 当地语言习惯与 SEO

多语种网站运营推广还需要熟悉当地语言和文化。如果只是简单地把中文版直译成外文，很可能看不到预期中的效果，处理不当还可能有反效果。

不同国家的购买习惯、文化、网站浏览习惯对网站转化的影响已经超出本书范围，就不去仔细讨论了。单纯从 SEO 角度出发，也需要考虑到语言、文化方面的不同，而这种不同对一个外语只是第二语种的人来说，要精确掌握并不容易。比如同样是英语，美式英语和英式英语就有不少区别。我们都知道很多单词美式和英式英文拼法不同，如颜色这个词，美式英文拼为 color，英式则拼为 colour。所有 SEO 的效果都体现在文字上，在进行关键词研究和网站内容的撰写上，都必须顾及到拼写的差异。

更不易掌握的是很多在中文中相同的事物，在英式和美式英文中是两个不同的词。如手电筒，美式英文称为 flashlight，而英式英文称为 torch。

反过来也一样，以英文为母语的人，做中文网站也不一定能选对关键词。我在一本翻译得很不错的 SEO 书中看到作者把 Anchor Text 翻译为“锚定文字”，其实在中文 SEO 和互联网行业，大家都称为“锚文字”。显然作者把 Anchor 这个词的原本机械行业中译法搬到 SEO 行业来了。如果这是一个美国 SEO 公司做的中文网站，选定的关键词与真正中文使用者的习惯用法相差就很大了。

简繁体中文也有类似情况。大陆把 SEO 翻译为“搜索引擎优化”，中国台湾称为“搜寻引擎最佳化”；大陆称为“网络营销”，中国台湾称为“网路行销”；大陆说“服务器”，中国台湾说“伺服器”；大陆称为“奔驰”的车，新加坡称为“马赛地”……

另外一个需要注意的是外部链接。不同语种的网站，最好建设相同语种的外部链接。经常看到有朋友问，自己做的英文网站建立的外部链接来自中文网站是否有效？这样的链接效果是有，但没有同样来自英文网站的外部链接好。

9.9 地理定位

很多做英文 SEO 的人经常会遇到这样的状况：SEO 人员在自己的电脑上搜索关键词排名不错，但客户在美国或欧洲搜索同样的关键词，排名却很差，甚至找不到。一般来说这就是地理定位（Geo-targeting）在起作用。

9.9.1 什么是地理定位

地理定位指的是搜索引擎根据用户所在位置及关键词本身的地理位置特性，返回不同结果。地理定位是个人化搜索的表现形式之一。相对于个性化搜索中的用户搜索历史、网站浏览历史等无法控制的因素，地理定位的影响因素中有一些站长可以自己控制或产生一些影响，所以单独拿出来进行比较详细的讨论。

地理定位在国际化的搜索引擎，如 Google、必应等表现得非常明显，不同国家的人搜索同一个词，看到的结果往往很不相同。

在同一个国家内，不同省份之间也经常会有地理定位的表现。如百度搜索的地理定位至少可以精确到省级。在百度搜索“SEO”，星箭的博客（www.starow.net ）就表现出明显地理定位特征。根据 4 月 20 号左右做的一个小调查（http://www.zaccode.com/space-2487-do-blog-id-3402.html），3 位北京用户搜索“SEO”时，星箭的博客都是排在第 3 位，而其他 12 位不在北京的用户搜索，星箭的博客都是排在第 8 位或第 9 位。显然，星箭的博客被百度认为带有北京地理特征，与事实相符。

9.9.2 地理定位的表现形式

地理定位有几种表现形式。

第一种最明显直观，搜索词中包含有地名，如“北京租车”和“上海租车”两个关键词，返回的结果肯定不同。虽然结果中的一些页面有可能是同一个公司的。

第二种是搜索引擎根据用户所在的地理位置以及用户账号中填写的住址等信息，返回不同搜索结果。比如北京的用户和上海的用户同样搜索“租车公司”，得到的结果通常也不一样。北京用户看到的结果以北京服务提供商为主，上海用户则看到上海服务商为主。英文网站也同样如此。在美国搜索“银行”和在新加坡或英国搜索“银行”看到的结果都不相同，虽然都是在 google.com 上搜索，但返回结果会偏重于用户所在国家的银行信息。

第三种表现形式是，用户使用同一个搜索引擎在不同国家的版本也会看到不同的信息。比如在 google.com 和 google.com.sg 和 google.co.uk 搜索同样的关键词，得到的结果也往往不相同。

做外文网站，更需要理解和关注地理定位效应。很多文字在不同国家使用，甚至是多个国家的官方语言。比如英文在英国、美国、加拿大、澳大利亚、新加坡都是官方语言。法文、西班牙文等也都有多个国家使用。不同国家的人使用相同的搜索引擎，搜索相同的关键词得到的结果可能有天壤之别。SEO 人员必须首先清楚自己的目标用户群在哪个国家，再通过下面提到的几个因素尽量影响自己网站在目标国家的排名，而不是自

己所看到的排名。一个中国站长做英文网站，自己看到的排名再好，英美国家的人看不到排名也没有任何效果。

9.9.3 地理定位的影响因素

影响地理定位的主要因素如下。

1．国家域名

不同国家有自己的国家域名。如果目标用户群限定在某个特定国家，那么网站最好就使用那个国家的域名。如针对英国的网站就使用.co.uk 域名，针对新加坡的就使用.com.sg 域名，针对国内用户的中文网站就使用.com.cn 或.cn 域名。具有国家后缀的域名是影响搜索结果地理定位的最重要因素。

2．主机 IP 地址

网站服务器最好放在目标用户群所在的国家。目标用户在英国就把主机放在英国，针对美国就把主机放在美国。

主机 IP 地址比国家域名作用小一点，但也可以辅助地理定位。比如.com 域名在全世界通用，虽然本来是用于美国网站的。在同样使用.com 域名的情况下，主机 IP 就成了判断网站地理属性的重要因素。.com 域名本身几乎很少能代表网站的地理属性。

3．本地商户登录

国际化的搜索引擎大多提供本地商户登录，在前面的地图搜索部分已经有所介绍。登录相应国家的本地商户信息，也是让搜索引擎判断网站地理属性的比较强的信号之一。

4．网管工具设置

Google 网管工具允许站长设置网站的地理位置。

5．联系地址及域名注册地址

网站页面上显示的公司联系地址、电话、区号、邮政编码及域名注册信息中显示的联系地址，都可以帮助搜索引擎判断网站与哪个地区相联系。所以在网站上的“联系我们”等页面上，应该清楚写明完整的联系地址，电话应该包括国家及地区区号。

6．本地网站链接

来自其他有地理属性网站的链接，也有助于自己网站的地理定位。尤其是一些具有鲜明地域特色的网上黄页、分类广告、生活资讯社区、本地商业组织、政府部门等类网站，都有着很明显的地理特征。来自这些网站的链接有助于提高网站地理定位排名。

7．页面语言

不同语言对网站的地理属性当然也有直接影响，比如中文往往显示目标用户是中国（但也不是百分之百）。有一些文字在多个国家使用，就需要有不同国家的语言特征，比如不同的拼法，同一件事情的不同称呼，都多少影响网站的地理定位信息。

9.10 社会化媒体的影响

社会化媒体网站是近几年互联网上最大的亮点。甚至有人说，在社会化媒体网站上搜索就快要取代在搜索引擎上搜索了，因为越来越多的人整天泡在社会化媒体网站上。这些用户的互动，亲戚朋友，喜怒哀乐，日常需求都被拴在了社会化媒体网站上。

社会化媒体是否能取代搜索引擎成为最主流的互联网商业应用还不好说，但不可否认的是，社会化媒体网站确实是当前互联网上最大的热门。数一数主流社会化网站，Facebook，Youtube，Twitter，维基百科，Digg，美味书签等，个个大名鼎鼎，流量和用户数巨大。全世界流量最大的 10 个网站中就有 4 个是社会化媒体网站。主要社会化媒体网站都有国内克隆版本。

社会化媒体网站已经影响互联网用户，尤其是新一代用户的生活方式和网站浏览方式，也不可避免地对网络营销及 SEO 产生不小影响。社会化媒体营销本身就是一种网络营销方式，近年来受到越来越多人的重视。

社会媒体营销所能带来的流量、关注度，经常具有爆发性，能迅速使品牌和产品信息在网上流传。可惜的是，到目前为止，社会化媒体营销有点叫好不叫座的味道，有流量有口碑，却不容易带来订单。从我自己的和所了解的其他一些网站所做的社会化媒体营销尝试看，转化率非常低。当然，这很可能是目前大家都还不完全了解社会化媒体营销的规律所导致的，还需要网络营销界不断探索。

总体上来说，社会化媒体营销值得大家关注，有时间和精力的网站也不妨进行尝试。

社会化媒体对 SEO 有一定的影响，甚至可以说这两者有不小程度的交叉。最欢迎社会化网络的人士常常有 SEO 背景。

社会化网络对 SEO 的主要影响表现在几个方面。

9.10.1 带来链接

社会化媒体网站上的活动产生在自己网站之外，对页面本身的优化没有影响，对 SEO 最重要的外部链接建设可能产生重大影响。

大部分社会化媒体网站是用户贡献内容，可以自由留下网址，但对 SEO 人员来说，

带来链接并不是发帖子留下网址这么简单。绝大部分用户能留下网址的地方都已经做了处理，留下的网址并不能成为普通意义上的链接。

比如 Facebook 上的很多内容需要用户登录才能看到，搜索引擎根本看不到用户写的东西，更不要说链接了。有的网站对链接审查十分严格，比如百度百科、百度知道。WebmasterWorld.com 完全不允许留任何链接，连网站的名字都不能提。点石互动对没有绝对必要的链接也一律删除。绝大部分搜索引擎能收录、用户又能留网址的地方或者导出链接使用脚本转向，或者做了 NoFollow 处理。也有的网站使用短 URL 服务，把用户留下的链接转换成短 URL，然后转向到目标网站。

可以说凡是用户能轻易留下网址的地方，社会化媒体网站都已经做了预防，不会让 SEO 留下正常链接。所以使用社会化媒体网站建设外部链接是间接的，需要在社会化媒体上让其他用户关注你的品牌、产品或话题，然后用户在自己的网站、博客上再次讨论和提到你的网站，从而带来链接。

正因为有了这样一个多余却必要的步骤，在社会化媒体网站上的普通互动很难带来链接。只有那些成为热门话题，比如上了 Digg 首页或在 Twitter 上被转发成千上万次，在论坛被热烈讨论的话题，才能带来大量流量及可贵的外部链接。通过这种方式带来的外部链接优势明显，非常自然，完全是用户主动贡献的，完全是单向链接，其中不乏高权重域名的链接。

9.10.2 互动及口碑传播

在社会化媒体网站上进行营销，最重要的是与其他用户互动，建立品牌知名度，有助于口碑传播。积累几个强有力的、有众多追随者的社会媒体网站账号，当网站有什么新鲜话题和产品时，一个微博或一句话帖子都能迅速传达到成千上万人，其中不乏在网上有影响力的人士，也不乏自己管理网站、博客的人士，他们的二次推动能使你的信息迅速传播开来。通过这种方式，你的文章和链接会出现在更多网站上，你的网站会有更多人讨论。由此带来的影响很难精确统计，长远来看却能影响排名和 SEO 效果。

9.10.3 新形式的链接流动成为排名信号

前面说过，留在社会化媒体网站上的链接，大多不是搜索引擎直接用来计算权重的正常链接。但是社会媒体网站用户数巨大，留下的链接数量也很惊人，俨然已经形成新形式的链接和权重流动模式。

普通网站很少会主动链接到其他网站。博客出现后多少改变了这种状况，博主一般还是很愿意链接到其他资源。随着 SEO 观念的普及，现在博客也惜链如金了。甚至有这样的趋势，原来积极写博客的人，近两年都不更新博客了，博客上的推荐和链接流动大大减少。原来写博客的人，现在反而更频繁地到微博和社交网站上推荐网址。以前提到某人时是链接到他的博客地址，现在提到某人时都改成链接到他的微博账号了。

大家可以想想，现在还能在什么普通网站上见到非交换的导出链接到其他网站？连主流门户转载文章都不给原出处任何链接了。再到 Twitter，Facebook 上看看，导出链接四处都是。社会化媒体网站上的链接、点击、用户流动已经是网上最活跃的一部分。

社会化媒体网站上的巨量链接流动，必然会被搜索引擎检测到，并且不太可能无动于衷。前面也讨论过，搜索引擎排名的目标是相关性、权威性和实用性。社会化媒体网站用户贴出来的链接就是实用性的一种表现形式。虽然这些链接噪声很高，不太可能像普通链接那样传递权重，但是很可能已经成为或将会成为新的排名信号，成为普通链接之外的辅助方式。

9.10.4 网络名誉管理

很多社会化媒体网站本身权重极高，网站上的页面排名能力也很强。这就使社会媒体营销成为网上名誉管理的有效方式，这一点在后面的 SEO 与品牌部分再详细探讨。

9.11 避免过度优化

过度优化这个概念有点悖论的意味。所谓“过度优化”，其实已经是错误优化，或者说不优化。所以严格来说不存在过度优化。但是为了讨论方便，我们也没必要咬文嚼字，还是使用“过度优化”这个用语。

SEO 人员应该了解网站上及网站外有哪些地方可以优化，应该怎样优化，同时也应该了解，当所有能优化的地方都被做到极致时，就可能产生负面效果，必须掌握一个度和平衡的问题。

如果一个网站具备下面一些特征，就可能被搜索引擎认为是过度优化，因而有被惩罚和降权的可能。

- 前面讨论的页面上可以针对关键词优化的地方，比如标题标签，H 标签，关键词和描述标签，黑体、斜体，内部链接锚文字，图片，ALT 属性，页面第一段文字，正文内容，URL，页面最后一段文字，这些应该突出关键词的地方都放上了关键词。
- 我们都知道外部链接锚文字非常重要，所以外部链接锚文字都使用目标关键词。
- 我们也知道来自高权重、高 PR 网站的链接很重要，所以想方设法获得的链接都来自高权重网站。
- 外部链接的锚文字与页面标题、H1 标签高度吻合。
- 外部链接很重要，所以 SEO 人员突击建设链接，短时间内外部链接快速增长。

当你把上面这些应该优化的地方都优化之后，就离过度优化的门槛不远了。如果你的网站在内容质量和用户体验方面确实在同行业网站中出类拔萃，那么问题应该不大。

但如果内容水平普通，甚至大部分内容是转载甚至抄袭的，只有 SEO 做得突出，那么过度优化就很可能是导致惩罚的最后一根稻草。

过度优化所导致的惩罚既可能是特定关键词排名下降，也可能是网站整体排名全部下降。如果你的网站排名下降很多，又找不到硬伤，过度优化就很可能是原因之一。

解决的方法就是“去优化”。网页上的标题、H 标签、链接锚文字等地方多一些变化，稀释关键词，使整个网站优化程度降到被惩罚的门槛之下。

网站内部优化就算做到极致，所能获得的排名分数也是有限的。而外部链接只要自然成长，潜力却是无限的。所以建议大家在做内部优化时，宁可欠缺一点，也不要全部做到极致。外部链接方面可以尽情发挥。把时间、精力放在创造高质量内容和吸引多种多样的链接上，比把时间花在页面细节上要有效得多，也安全得多。

9.12 SEO 与品牌

在大部分情况下，SEO 的目标是带来搜索流量和直接销售。同时，SEO 对企业品牌建设也起到一定的作用。

9.12.1 排名第一就是品牌

普通用户往往认为搜索行业主要关键词时，排在第一页的网站是比较可靠的品牌和服务商，如果能排在第一位，就更被认为是可以被信赖的。当然这个假设并不是百分之百正确，不过现实如此，互联网用户确实有这种倾向。一些被误导的例子反倒更能证明这一点。前几天刚刚看到报道，一个用户搜索“机票”，看到排在第一位的网站后，没有任何怀疑就认为那应该是航空公司的官方网站，直接付钱订票，结果被骗，排在第一位的是一个假冒的骗子网站。

主要行业关键词流量不一定很大，排在前面的品牌价值很高。用户搜索行业主要关键词时，哪个网站排在第一，就会被很多人下意识地认为是最好的品牌。这与流量没有关系，用户无论是否点击，是否真正访问你的网站，只要看到在搜索结果页面中你的网站排在第一，都会造成这种印象。

我在新加坡的主业是提供服务器和虚拟主机服务，相关关键词数年来一直在 Google 排名第一。收到过来自本地银行的询问，要把银行的数据库放在我们的服务器上。还有用户在沟通时告诉我们他是律师，如果我们想上市他可以帮忙。对他们来说主要关键词排在第一位，就意味着是个大公司，是个著名品牌，虽然这往往不一定是正确的——我们只是个小作坊。

用户搜索公司品牌时，真正的官方网站排在第一是必需的，应该不惜任何代价做到。幸运的是，这往往并不困难。

9.12.2 传统品牌建设与 SEO 结合

公司名称、口号、广告语等品牌元素应该与 SEO 完美配合。也就是说搜索特定的公司名称、口号、广告语时，公司官方网站必须排名第一，不仅加强品牌，也可以避免造成不必要的误导。欧美大公司对这一点都很看重。举个简单的最新例子，电影“2012”推出后，在 Google 搜索“2012”排在第一的就是这部电影的官方网站，如图 9-6 所示。

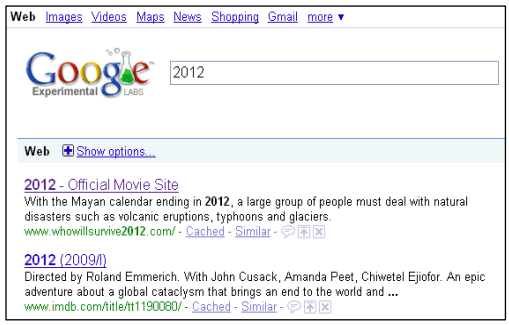


图 9-6　在 Google 搜索“2012”

国内很多企业在这方面认识不足，搜索特定产品牌子、广告语，公司或产品官方网站完全找不到。

9.12.3 用 SEO 进行网上危机公关

网络名誉管理也是 SEO 可以发挥作用的地方。网络危机公关和名誉管理是近两年非常热门的话题。网络上的危机事件常常起源于论坛，爆发于搜索引擎，所以搜索引擎、SEO 与网络危机公关有着天然的联系。从事 SEO 的人现在都意识到控制负面新闻、进行名誉管理及危机处理是 SEO 的强项之一。

利用 SEO 进行网络危机公关，原理很简单：搜索自己品牌或公司名称出现负面新闻、评论时，把自己控制的网站页面通过 SEO 做到前面，把负面新闻挤出前两页或前三页，但执行起来却不简单。

要想有效压制负面新闻，通常需要占领前两页到前三页的搜索引擎结果。这 20～30 个网页靠自己的网站很难实现，无法快速反应。当网络危机和负面新闻出现时，建设 30 个网站进行 SEO，等这些域名有了权威度，排名上去时，新闻事件早已成昨日黄花。

利用企业官方网站也很难。同一个网站上的不同页面通常不会多次出现在搜索结果前面，能出现两次就算不错了。Google 甚至对同一个主域名下的二级域名也这样处理。对百度来说，一个域名下建多个二级域名，目前还可以实现多个排名，但是往往有作弊之嫌，很可能得不偿失。

所以靠自己的网站做网络危机公关不是一个有效的方法。动用公关部门，联系论坛、博客删帖也不一定能达到效果。能联系几十个大的论坛删帖，但很难联系几千、几万个转载的小论坛，删不胜删。

仔细观察一下搜索引擎返回的负面信息就会发现，大部分这类新闻或评论来自于已经有权威度的论坛、新闻门户、书签网站、博客等地方。这也是运用 SEO 进行网络危机公关与社会化媒体网站有紧密联系的原因，因为排在前面的有很多是社会化媒体网站上的负面消息。

要把这些负面新闻挤出前两页，最好的方法就是在这些已经排到前面的论坛、博客、新闻门户、书签、视频等网站上发表正面帖子或正面新闻。比如你看到一篇负面新闻来自某书签网站，就在这同一个书签网站上提交一个正面消息网页。排在前面的其他论坛、博客、新闻门户等也同样处理。当然，除了已经排到前面的网站，到更多网站发帖效果更好、更快。

你所提交或发表的内容，与那些已经有好的排名的负面新闻具有同样的域名权重（来自同一个域名），排名能力不会输给那些负面新闻，你的正面信息已经有 50%左右的机会超过负面消息。唯一的区别只是你发表的内容可能时间要晚一点。这时可以再进一步，给这些你提交或发表的正面内容造几个外部链接。不需要很多，几个链接就足以使这些正面新闻页面权重超过那些负面新闻，因为很少有发表负面评论的人还给自己的评论造外部链接。

这样，你的正面信息具有同样的域名权重、同样的相关性及更多的外部链接，足以把负面信息挤出前两页或前三页。

如果这些负面信息页面竟然还有很多外部链接，这说明你碰到的是个行家，不是普通消费者在发牢骚，或者发牢骚的是位权威人士，内容被更多人转载和链接，那么你需要更高级的策略，做更多弥补工作。

在没有危机事件发生时，企业也应该尽量在常见社会化媒体网站建立账号，发布企业信息。用户搜索与企业有关的品牌名称时，如果返回的主要是企业控制下的社会化媒体网站内容，就对日常推广和危机控管都做好了准备。

做好网上名誉管理还要有敏锐的网上信息搜索能力，在负面新闻刚出现时就能发现并及时处理。在这方面 Google Alerts 是一个非常好用的工具。用户在 Google Alerts 设定要监测的关键词，通常是公司名、品牌名、老板名字或者你自己的名字，一旦设定关键词出现在网络上，Google 就会通知你。用户可以选择 E-mail 通知或者 RSS 订阅，可以选择监测博客、视频、新闻、普通网页，也可以这些同时监测。

虽然 SEO 是网络危机控管的有力手段，但毕竟不是治本的办法。即使不算马后炮，也只能算亡羊补牢。网络危机公关和名誉管理远远超出 SEO 的范围。危机和品牌管理要

想做得最好，关键不在危机与品牌管理本身，而在于产品、服务的质量及售后支援。满足了客户需求，没有被愚弄、被欺骗、被激怒的用户，自然就没有负面评论，这才是根本。

9.13 针对不同搜索引擎的优化

经常有人在博客和论坛上问：Google 和百度的优化有什么不同？简单说，针对不同搜索引擎的优化，原理相同，方法也相同，只是在一些细节上有不同而已。长远来说，不同搜索引擎的未来趋势是越来越相似，SEO 人员应该更多地考虑搜索引擎的相似之处。

9.13.1 SEO 原则不变

六七年前，SEO 行业曾经有过这样的阶段，针对不同搜索引擎做不同的网站，使用不同的优化手法。比如针对 Google 做一个网站，迎合 Google 的口味，禁止雅虎、AltaVista 等其他搜索引擎收录。再做另外一个网站去迎合 AltaVista 胃口，同样禁止 Google 和其他搜索引擎收录。现在已经没有人再这样做了。一方面是因为 Google 占市场份额的 70%～80%，只要迎合 Google 就足以了。另外一方面也是因为 Google、雅虎、微软，包括百度的排名算法越来越趋同，完全没有必要做不同的网站，使用不同的方法来优化。

从实践上来看，只要抓住 SEO 的根本原理，做好基础优化，一般来说在所有搜索引擎中排名都会不错。不同搜索引擎算法上的细微差别造成排名不同是正常的。但对一个真正优秀的网站来说，不会有天壤之别。

9.13.2 百度和 Google 的区别

作为了解内容，下面举几个 Google 和百度的不同之处。

在收录方面，Google 很容易收录新网站和新页面，一两个质量不必很高的外部链接，就能让新网站收录。网站上有转载内容影响也不大。不过 Google 有沙盒效应，收录门槛很低，获得好的排名比较难。百度正相反，很多新网站要被百度收录是个难题，有时需要比较长的考核期。而一旦被收录，比较容易获得排名和流量。网站刚开始时的原创性对百度来说比较重要。

百度有比较大的首页优势，而 Google 对所有页面一视同仁，不管是首页还是内页。在搜索结果页面上，Google 较多返回网站内页，而要想在百度上获得较好的排名，经常需要靠首页。这就可能影响网站架构及不同关键词在网站上的分布处理。

Google 对外部链接很重视，对页面元素则没有那么敏感。我们经常可以看到排在第一页的页面，关键词在页面上出现一次的也有，出现十次的也有。Google 对于关键词在页面上的出现次数、位置，看不出明显的、有规律的偏好。而查看排在前几十位的页面的外部链接，就能看出比较一致的趋势，排在前面的通常比排在后面的外部链接质量高，数量也多。

百度对外部链接的依赖性比较小，对页面本身的相关性却比较敏感。这种敏感既体现在，在正确的地方出现关键词有助于排名，也体现在关键词显得堆积时更容易引来惩罚。相比之下，Google 既不会因为关键词出现次数多给予好的排名，也不会因为关键词出现次数多给予惩罚。

百度排名经常有大起大落。我接触的一些网站常常在优化后的某一个特定时间，很多关键词排名同时上升，流量也突然增长。而 Google 对网站的排名处理总是渐进式的，在几个月的时间内平稳增长，没有一个时间点上有爆发式增长。对网站的惩罚同样如此，百度上的排名可能在一夜之间全部消失，Google 则较少出现这种情况，除非严重作弊。

百度对网站更新比较看重，持续有规律地增加内容往往可以有效地提高在百度的排名。Google 对内容更新没有百度那么敏感。当然持续有新内容肯定是件好事，但很多网站几年没有任何更新，排名也不会下降。例外的是对一些有时效特点的网站，比如博客和新闻网站，Google 也要求持续更新。

总的来说，上面这些不同并不是本质上的区别。SEO 人员只要做好基本优化工作就可以了，不必拘泥于细节上的不同。

9.13.3 英文网站优化

这里说一下与不同搜索引擎优化相关的话题，英文网站的优化。常有站长问我怎么优化英文网站，与中文网站有什么区别。简单说，优化英文网站与优化中文网站没有本质区别，从关键词调查，到网站架构，再到页面优化，以及外部链接建设，整个过程和方法是一样的。

由于中文网站优化应该更关注百度，英文网站则自然以 Google 为目标，中英文网站优化的区别很大程度上就是百度和 Google 算法上的不同。大家可以上前一节和排名因素调查一章看到百度和 Google 给予排名因素的不同权重。

优化英文网站最重要的倒不是过多考虑 SEO 方面的不同，而是过英语关。语言不过关，可能从关键词研究阶段就错下去了。

做英文网站也应该考虑欧美用户的网站使用习惯。比如，欧美用户更习惯简洁的页面设计，中文门户那种满满长长的、广告四处闪，甚至四处飞的首页，搬到英文网站上一定不会受欢迎。英文用户点击链接习惯于在原窗口打开，很少有中文网站这样在新窗口打开页面。

英文网站建议放在美国或欧洲的服务器上。国内服务器有的在国外打开速度还不错，有的极慢，严重影响用户体验。

9.14 网站改版

这里所说的网站改版指的是网站主题没有变化，页面设计或网站架构做出比较大的改变。如果网站主题内容发生重大改变，把域名从一个行业换成另一个行业的内容，从SEO 角度说，这不是网站改版，而是网站自杀。

首先应该明确，网站改版是万不得已时才做的变化，能不改最好不要改，尤其是网站结构。不过很多时候由于种种原因，如需要改进用户体验或公司发生了合并收购等变化，设计风格也需要变化，甚至有时候公司换了个老板，都可能不得不改版网站。

9.14.1 设计还是 CMS 系统改变

进行网站改版时，首先要确定网站只需要页面设计的改变，还是需要改用新的 CMS 系统。如果只是页面设计方面的改变，通常页面内容不会有什么大的变化。很多内容与表现分开的 CMS 系统，更换或修改模板，页面 HTML 代码几乎没有什么变化，页面视觉展现却可以完全不一样。这种情况对 SEO 几乎没有什么影响，可以放心进行。

如果需要新的 CMS 系统，就要小心计划和执行。像前面所说的，能不换尽量不换，如果必须要换，尽量做到网站 URL 命名系统不要更改。如果旧的 CMS 系统目录及文件名命名规则，使用 URL 重写模块，可以在更换 CMS 系统后还保持原来的 URL 结构。

9.14.2 不要改 URL

不到万不得已，千万不要更改 URL，否则对搜索引擎来说，新的 URL 就是新的页面。整站 URL 改变，几乎意味着这是一个新的网站，搜索引擎需要重新收录所有页面，重新计算排名。同时旧的 URL 还在搜索引擎数据库中，势必会造成复制内容、页面权重分布的混乱，网站排名及流量都会有一段混乱期。对大的网站来说，整个网站重新爬行抓取一遍可能需要几个月的时间。

如果实在没办法保留 URL，那么尽量从旧 URL 做 301 转向到新 URL。如果连 301 转向都不能做完整，至少也要挑选重要页面做 301 转向。这里所说的重要页面包括栏目首页、带来比较多搜索流量的内容页面、有比较多外部链接的页面。如果不能使用 URL 重写模块及正则表达式进行整站 301 转向，至少也要人工挑选出这些重要页面做 301 转向。

9.14.3 分步更改

网站改版时尽量不要同时更改导航系统。对主要导航系统的修改，往往会使网站上链接结构、页面权重的流动和分配产生重大改变，处理不好将会对网站新页面的收录造成影响。所以应该在网站 CMS 系统或 URL 系统修改完成几个月之后，收录已经恢复原有水平之后再修改导航系统。

同样，无论是页面设计改版，还是采用新的 CMS 或 URL 系统，不要同时修改网站内容，要确认网站收录没有问题之后，再优化页面内容。

对大的网站来说，无论怎样小心，重大改版经常不能保证顾及到所有 URL，这时一个恰当的 404 页面就变得很重要了。404 页面的设置请参考页面优化部分的说明。正确设置 404 页面，搜索引擎会自动把已经不存在的页面从数据库中清除，但需要一段时间。

网站改版经常是商业决定，不是 SEO 部门所能控制的。SEO 人员必须参与到改版的计划过程中，提前认真规划，预想到所有可能的情况，尽量一次改版正确，避免改来改去。尤其是 URL 有变化时，搜索引擎对大量新出现的 URL 会很敏感，需要重新收录、计算权重，也很可能出现一段沙盒期。如果不能一次完成，多次更改很可能对网站产生重大负面影响。

9.15 Google Dance

Google Dance 是 SEO 行业很有名，但现在经常被误解的一个概念。

9.15.1 什么是 Google Dance

Google Dance 指的是 2003 年以前，Google 每个月进行一次整个数据库和算法更新。在这个更新过程中，页面索引库加入新页面，删除旧页面，排名算法也有重大改变，很多网站排名有大幅变化。由于 Google 有很多数据中心，新的索引库及算法要在多个数据中心同步更新需要一段时间。在这段 Dance 期间，一部分网站排名上下浮动，不同用户看到的排名也不一样，同一个用户在不同时间可能访问的是不同的数据中心，看到的排名也可能不一样。所以整个 Google 搜索排名处于一种上窜下跳的状态，这也就是叫“Google Dance”的原因所在。

2002 年 2 月开始，网管世界（WebmasterWorld.com）给这个每月一次的 Google Dance 进行了命名，一般是以台风的名字命名。这些 Google Dance 的名称也得到了 Google 官方的认可。比如 2003 年的几次 Google Dance 名称如表 9-1 所示。

表 9-1　2003 年的几次 Google Dance 名称

时　　间	Google Dance 名称
2003 年 2 月	Boston
2003 年 4 月	Cassandra
2003 年 5 月	Dominic
2003 年 6 月	Esmeralda
2003 年 11 月	Florida

2003 年 11 月进行的佛罗里达（Florida）更新，是所有 Google Dance 中最著名的一次，SEO 从业人员也都耳熟能详。在这次更新中，很多经过 SEO 的网站排名一夜之间消失得无影无踪，一些网站排名从此再也没回来过。佛罗里达更新对 SEO 行业产生了地震式的影响。

9.15.2 Google 已不再 Dance

2004 年以后，Google 已经停止了 Google Dance。这是被国内很多 SEO 人员误解的地方。直到现在很多人一看到网站的 Google 排名有些变化，就说 Google 又 Dance 了，其实 Google 已经不 Dance 很久了。2004 年以后，Google 排名更新都是以一种不间断的，但是小幅度的方式进行，不是每个月大更新一次，新的网页不断进入数据库，新的排名算法也随时启用。据 Google 自己表明，一年之中 Google 会有几百次算法改变。这种新的更新方式被称为 everflux。

现在国内 SEO 行业的人大多是 2003 年以后才开始接触 SEO 的，有六七年经验的应该算很资深了。没亲身体验当年 Google Dance 的人很难对此有感性认识，不能感受到 Google Dance 与现在 everflux 更新的区别，只是一看到排名变化就认为是 Dance 了。

9.15.3 近年 Google 更新

2004 年之后，还有几次比较重大的 Google 排名变化，网管世界也给予了命名。不过这些更新都不是 2003 年以前传统意义上的月度 Dance。对于 2004 年之后网管世界的命名，Google 的态度都是：好吧，既然你们给了名字，就这么叫吧，但其实不是一次 Dance。

2004 年后几次比较有名的更新是：

- 2005 年 9 月，Jagger 更新。Jagger 是所谓“Google 大爸爸数据中心”（Big Daddy）中的一个准备阶段。“大爸爸”是 Google 对自身整体架构的一次全面更新。
- 2009 年 2 月，Vince 更新。在这次更新中，一些大品牌网站排名显著上升。Google 对这次更新的说明是，这是由一位叫 Vince 的员工所写的排名算法中的一个小更改造成的，这就是该名称的由来，但其实没有专门针对大品牌做什么改动。
- 2009 年 8 月，Caffeine（咖啡因）更新。所谓咖啡因是 Google 对基础架构的一次重写，尤其是索引部分的代码。经过重写的系统扩展性更强，速度更快。咖啡因在 2009 年年底之前只在一个数据中心启用，2010 年之后才在其他数据中心陆续启用，目的是为了避免像 2003 年年底的佛罗里达更新那样，在圣诞购物热潮期间对太多网站造成重大影响，甚至有的公司因为佛罗里达更新而倒闭了。
- 2011 年 2 月，Panda（熊猫）更新。这次算法变化主要目的是减少搜索结果中低质量页面的数量，大量网站受到影响。名称和 Vince 更新一样，源于相关算法修

改主要工程师的名字叫 Panda。2011 年全年还不定期推出 Panda 更新算法的微调，被 SEO 界称为 Panda2.0、Panda3.0 等。

Google Dance 的另外一个意思是指 Google 在 SES 搜索引擎大会上组织和赞助的晚会。这个被称为 Google Dance 的晚会也已经停止了。

9.16 Google 全站链接

Google 全站链接官方名称是网站链接（Sitelinks），中文 SEO 行业一般称其为全站链接。

9.16.1 全站链接的出现

Google 全站链接指的是排名列表的一种扩展显示，某些权重高的网站除了正常搜索结果外，还会显示最多八个四行两列内页链接，如图 9-7 所示。

图 9-7 Google 全站链接

全站链接给网站更大篇幅，也提供了更多点击访问入口，能在很大程度上提高点击率。数据表明，出现全站链接时，比原来的普通排名点击增长 40%。

全站链接的出现和链接选择是完全自动的，由 Google 算法生成，站长本身并不能控制。影响全站链接的可能因素包括以下几点：

- 全站链接只出现在权重高的域名上，很少有年龄两年以下的网站出现全站链接。这也使得全站链接成为判断网站权重的一个不错方式。
- 全站链接只有在网站长期稳定排在第一位时出现。也就是说，只有网站是某个关键词的最权威结果情况下才会出现。搜索品牌或网站名称是最常出现全站链接的。
- 有全站链接的网站并不是搜索所有关键词时都会出现，只是几个最相关、排名最好的关键词才出现。而且这几个关键词都需要达到一定的搜索量。一些长尾关键词就算网站排名第一，通常也不会出现全站链接。
- 出现全站链接的网站需要具备良好的导航及内部结构。Google 算法是通过网站链接结构选择出最多 8 个内部页面。一般来说这 8 个内部页面都在首页上有链接。
- 除了链接结构，流量分布也很可能是因素之一。也就是说选择出的 8 个内页，通常是用户最经常点击访问的页面。

- 内部链接锚文字对全站链接也有很大影响，尤其是链接标题经常是由内部链接锚文字决定的。页面本身的标题标签、H1 标签也必须有高度相关性，对全站链接标题也有一定影响。

9.16.2 屏蔽全站链接

注册了 Google 网管工具的站长，可以在网管工具后台查看所生成的全站链接，还可以屏蔽某些内页，如图 9-8 所示。

图 9-8　在 Google 网管工具屏蔽全站链接

一些不太重要的页面被 Google 选定为全站链接时，站长可以将这些链接屏蔽掉。在网管工具中单击“拦截”按钮，就可以将对应的链接删除。不过使用这个功能一定要非常小心，因为屏蔽一个链接，并不意味着 Google 会自动加入另一个链接补齐 8 个。Google 算法会重新计算其他内页，只有在其他内页也符合条件时才会加进来。所以只要生成的内部链接不太离谱，轻易不要屏蔽，不然可能只剩下少于 8 个全站链接。在网管工具后台屏蔽全站链接，只有 90 天有效期。

9.16.3 迷你全站链接

除了四行两列的经典全站链接外，Google 还有一个所谓的迷你全站链接。官方名称是单行全站链接（Oneline Sitelinks），也就是只显示一行最多 4 个内页链接，如图 9-9 所示。

与经典全站链接不同的是，迷你全站链接不要求网站是针对特定关键词的最权威网站，只要内容相关就有可能出现。所以迷你全站链接既可能出现在排名第一的网站，也可能出现在排在后面的网站。在同一个搜索结果页面上，还可能出现多个网站都有迷你全站链接的情形。

迷你全站链接的 4 个内页通常就是经典全站链接 8 个页面中的前 4 个。

SEO每天一贴 – Zac的SEO博客，正在写书，暂时每月两三贴
SEO每天一贴研究搜索引擎优化SEO技术，网络营销及电子商务思考。... 介绍和研究搜索引擎优化SEO技术，Zac的SEO优化及网络营销，电子商务，网站推广运营思考。...
网络赚钱 - 我的书 - 英文网站SEO - SEO技术
www.chinamyhosting.com/seoblog/ - 9 小时前 - 网页快照 - 类似结果

图 9-9　迷你全站链接

9.17　个人化搜索

个人化搜索也是 Google 最先开始尝试、其他主流搜索引擎也跟进的排名方式。所谓个人化搜索，指的是不同用户搜索相同关键词，看到的排名不相同。

9.17.1　什么是个人化搜索

广义的个人化搜索包含三方面意义：

- 一是地理定位，来自不同地区的用户看到的搜索结果不同。
- 二是社会搜索，有不同社会化网络联系的人，会看到不同的社会化网络内容。
- 三是根据用户的历史搜索数据返回不同结果。

地理定位前面已经讨论过。社会搜索指的是用户可以在 Google 账号中添加自己注册过的其他社会化媒体服务如 Facebook、Twitter 等，Google 在搜索结果页面中会显示来自社会化媒体网站的你的好友发布的相关内容。

本节只讨论狭义的源自于历史搜索数据的个人化搜索。

2007 年 Google 就开始根据用户搜索历史、点击、访问数据，提供不同的、更适合特定用户的排名结果。比如说用户搜索“图书”这个关键词，然后经常点击访问当当网，而不是其他网上书店，这样的搜索历史会被 Google 记录下来，久而久之，Google 就会在用户搜索与图书相关的词时，更多地返回当当网的页面。

再比如，用户近期经常搜索与旅游相关的信息，当他搜索“巴黎”这个词时，返回的结果将更偏重于巴黎旅游信息，而不是巴黎的香水、化妆品、时装信息。

基于历史搜索数据的个人化搜索刚推出时，只限于登录了 Google 账号的用户。用户可以选择关闭历史搜索数据，不使用个人化搜索。2009 年 12 月，Google 开始把个人化搜索应用于所有用户，包括登录了 Google 账号和没有登录 Google 账号的。所以一个用户在使用同一台电脑时，就有可能出现两套个人化搜索数据。登录 Google 账号时，Google 会永久性地保存与这个账号相联系的所有历史数据。退出 Google 账号后，Google 还会通过 Cookie 记录这个用户的历史搜索数据。不过这时所记录的与登录状态下的是两套不同的数据。未登录 Google 账号用户的历史数据只保留 180 天，用户删除 Cookie，这些数据也就消失了。

9.17.2 个人化搜索对 SEO 的影响

个人化搜索是搜索引擎未来发展的必然趋势之一，对于理解用户搜索意图、进而返回更适合的内容有很大帮助。如果个人化搜索更全面使用、造成的排名结果差异也更大，传统意义上的关键词排名将失去意义，因为不同用户看到的搜索结果很可能完全不同，一个关键词对用户 A 排名第一的网站，对用户 B 却可能不在前两页，甚至更靠后。这无疑将让传统意义上的 SEO 消失于无形。

从这个角度看，未来的 SEO 必须更注重用户体验和黏度。只有用户认可、愿意多次访问的网站，才能保持在更多用户搜索时出现好的排名。

域名权重将更为重要。用户第一次搜索特定关键词时，大网站、权重高的网站，才有机会出现在第一页，才有机会让用户点击访问，进入历史搜索记录，从而使更多关键词排名更上一层楼。大网站、权重高的网站也往往有海量页面，大大提高众多用户发现这个网站的机率，产生前几次点击访问，进入历史搜索数据。强者更强，小网站连第一次被发现的机会都很小，就更不用说进入用户的个人化搜索了。

所幸，就目前来说，个人化搜索的影响还不大，尤其是前三名结果，绝大部分用户无论登录 Google 账号与否，看到的排在最前面的结果基本变化不大。未来个人化搜索是否会明显改变大量关键词的前三个结果，值得 SEO 人员关注。

第 10 章
SEO 观念及原则

前面的章节讨论了具体 SEO 技术和细节。完全聚焦于技术容易只见树木，不见森林，本章我们就把镜头拉远一点，从总体上审视 SEO，谈谈 SEO 观念和原则。

10.1 搜索引擎的目标

搜索引擎的目标到底是什么？看似简单的问题，很多人不一定能答对。要深入理解 SEO，就要深入理解搜索引擎本身的目标是什么。

10.1.1 搜索引擎的目标是满足搜索用户

用搜索引擎自己的话来说，百度的使命是“让人们更便捷地获取信息，找到所求”。Google 的使命是“整合全球信息，使人人皆可访问并从中受益”。

搜索引擎自己表明的使命写得比较宏大，其实简单说就是：用户来搜索任何关键词时，都能找到需要的信息。

搜索引擎的用户是网上搜索信息的人，客户是广告商。站长们（以站长身份出现时，而不是搜索用户）说到底不是搜索引擎的用户，更不是他们的客户。搜索引擎并不欠站长或 SEO 人员什么东西，网站收录不收录，排名怎么样，都是搜索引擎自己的事。就算我们的网站被完全删除，其实也没什么好抱怨的。

目前所有的搜索引擎都是通过搜索竞价广告赢利的，不同搜索引擎的区别只在于竞价广告出现的位置、数目及是否比较清晰地标明是广告，其 PPC 本质是一样的。要想通过搜索广告赢利，就必须有搜索用户使用搜索引擎。

更换搜索引擎服务商的成本近乎为零，这是搜索引擎最大的风险之一。我们使用的其他互联网服务，想更换或多或少都有些麻烦，比如换 E-mail 地址，把博客从一个提供商搬到另外一个提供商，从一个 SNS 阵地换到另外一个等。这些都可以做到，只是需要些时间和工作。人都是很懒的，凡是需要花时间精力的，能避免则避免。更换搜索引擎是成本最低的，从使用百度换到使用 Google，或者反过来，用户既不用费时间，也不用花钱，也不费事。

这就决定了搜索引擎要想保持甚至提高搜索份额，进而通过广告赢利，就必须最大程度地满足用户搜索需求，也就是返回让用户满意的信息。搜索引擎不断推出新产品、更新算法、更新数据库，所有的工作都是围绕着返回相关信息这个根本点。失去这一条就失去用户，就失去赢利。

10.1.2 搜索引擎不在乎我们

深入了解这一点，对 SEO 思维其实有很大影响。看似简单的道理，使很多 SEO 迷思可以迎刃而解。

举个例子。有的站长抱怨自己费了很多心思去做网站，没有作弊，信息很全，为什么排名就上不去呢？甚至为什么搜索引擎就不收录呢？我们站在搜索引擎的角度去想就会明白，对搜索引擎来说，收不收录你的网站，是否给予排名，搜索引擎一点都不在乎，只要搜索用户能找到他们需要的信息就足够了。至于信息来自哪个网站，搜索引擎本身是无所谓的。

搜索引擎缺了哪个网站都没关系，除非你的网站是搜狐、新浪级别的。如果在搜索结果中找不到新浪搜狐这样的网站，那么搜索引擎的用户体验会大大降低。但是对于绝大多数中小企业和个人网站来说，缺了我们，搜索引擎还真无所谓，因为搜索用户无所谓，他们本来也没期望看见我们的网站。

一个网站是站长的心血，但是同样质量、同样信息的网站，没有几十万个恐怕也有几千几万个。缺了哪个网站，对搜索引擎、对搜索用户都没什么影响。

以 Google 为代表的部分搜索引擎一直希望与站长社区沟通，以方便搜索引擎更快、更好、更全面地收录网上信息。这是就整体而言。没有站长们的网站，搜索引擎就没有信息，就无法满足用户。但对特定网站来说，搜索引擎就无所谓了。

这也就是为什么新闻集团的默克多一直抱怨 Google 在自己的搜索结果中显示新闻网站内容，赚取广告费，新闻网站本身却没得到什么（我们姑且不论这个观点是否符合事实），Google 的回应很明确，如果新闻网站不想被显示在搜索结果中，请放上 Robots 文件禁止收录，一切就都解决了，Google 既不会收录新闻集团的新闻网站也不会显示。Google 无所谓。没有了新闻集团的网站，还有成千上万其他新闻网站哭着喊着要被收录、被显示呢。

同样，淘宝禁止百度收录，百度也不太有所谓。网上还有无数电子商务网站的信息能满足搜索用户的要求，对百度信息质量并不会产生实质影响。

明白这一点，站长的心态才能平和。有了你和缺了你，都不影响搜索引擎的目标。

10.1.3 搜索引擎在乎垃圾

再比如对于垃圾网站的打击上，SEO 人员也往往缺乏足够的认识。让我们再站在搜索引擎的立场去考虑。返回垃圾内容，尤其是挂羊头卖狗肉的无关内容，对搜索引擎用

户体验是个很大的伤害。如果我们在一个搜索引擎搜索减肥方法，看到的净是六合彩信息，一次两次可以容忍，老是这样我们就再也不会用这个搜索引擎了。

所以千万不要小看了搜索引擎无情、严厉打击垃圾网站的信心和动力，垃圾网站其实是严重影响搜索引擎赢利能力的。

从另一个方向看，有时候 SEO 人员也会抱怨某某网站在作弊，可是却没有被惩罚或删除，大叹不公平。这也是没有深入理解搜索引擎的目标。搜索引擎对特定网站既不偏爱，也没有深仇大恨，只要做到把垃圾内容控制在一定比例之下就足够了。搜索引擎并不会花太多的人力物力，针对特定网站必须要怎么样，而是通过算法把大部分垃圾清除，让用户很少看到不相关内容，就已经完成目标。

做网站和做 SEO 的人现在都知道，首先要从用户出发，做用户喜欢的，搜索引擎就会喜欢。同样，遇到 SEO 方面的疑问时，也要站在搜索引擎的角度思考。你网站上的内容或方法是否与搜索引擎本身的目标相冲突？还是没什么影响？诚实回答了这个问题，一些疑惑就会迎刃而解了。

10.2 相关性、权威性、实用性

大部分关于 SEO 的文章喜欢探讨细节问题，包括我自己的 SEO 博客。探讨细节问题，容易让初学者知道从哪里下手来优化网站，所以更受站长欢迎。当你掌握了技术细节后，还需要跳出来，从宏观上看，到底什么样的网站在搜索引擎排名中有优势。

在我看来，一个网站要想被搜索引擎喜欢，必须要有相关性、权威性和实用性。

10.2.1 网站内容的相关性

也就是用户搜索的关键词与网页内容是否匹配，是否有相关性。

相关性的加强可以通过页面内优化和一部分链接优化来达到，包括网页标题、页面内的关键词位置布局、关键词的强调、写作时考虑语义分析、内部链接的安排，以及外部链接的锚文字、链接页的内容、链接源网站的主题等。

内容的相关性是做网站的人最容易控制的，也是最容易被作弊的。第一代搜索引擎主要以相关性做排名判断，但在被钻空子钻得一塌糊涂后，不得不引入权威性衡量。

10.2.2 网站及网页的权威性

网站或网页的权威性大部分是通过外部链接来衡量。高质量的外部链接越多，网站或网页本身的权威性就越高。另外，域名注册历史、网站的稳定性、是否有隐私权政策、联系地址等一些细节，也会在一定程度上影响网站的权威性。

另外要注意的是，外部链接对网站权威性的影响是有选择性的，也就是说来自相关内容网站的链接对提高权威性帮助最大，不相关内容的链接帮助很小。比如我要是在SEO每天一贴首页加一个链接到某个美食网站，对对方的权威性几乎没什么帮助，因为很明显，我的博客本身在美食方面就没有任何权威性。

网站的权威性不能被站长自己完全控制，要想作弊比较费时费力，大量群发现在也越来越容易被检测。但是在某种程度上，权威性还是可以被操作。无论是花钱还是花时间，都可以得到更多的人为链接，现在搜索引擎开始考虑网站的实用性。

10.2.3 网站的实用性

也就是对用户来说，你的网站到底有多大用处？用户是不是喜欢你的网站？

如果用户在你的网站花的时间多，浏览页数多，三天两头经常来看你的网站，还加入了书签，并且在不同的社交网站上评论，这些都可以帮助搜索引擎理解你的网站对用户的实用价值。

搜索引擎的工具条、广告代码、流量统计服务等可以帮助收集这类信息，越来越多的社会媒体网站也能表现网站的受欢迎程度。

网站的实用性想作弊就更难了，因为你没办法控制用户电脑和用户的行为方式。当然并不是完全没有可能。不过如果你的网站在相关性、权威性、实用性上都很出色，还都是作弊得来的，这可能性就很低了。

10.3 SEO与赚钱

通过SEO赚钱是所有个人站长都关心的问题，本节就简单探讨一下，对商业化的公司网站也有一些借鉴意义。

通过SEO赚钱无非有两个出路：一是帮别人做SEO，二是给自己做SEO。

10.3.1 给别人做SEO

先说帮别人做SEO，也就是某种形式的SEO服务。

1．SEO服务的优势

提供SEO服务有很明显的优势：入门快，成本低，市场大。网上有无数关于SEO的博客、论坛、电子书，您现在手里就正拿着一本最好的SEO入门教程。只要有一定的网站制作基础，把教程读完之后，做一两个网站实践一下，对SEO的过程和效果有了一定认识，就可以提供SEO服务了，整个学习过程几个月就可以完成。相对于许多其他行业，SEO入门算很快速了。

提供 SEO 服务也不需要任何设备，只要有一台能上网的电脑足矣。刚开始时连办公室也不用，自己在家里就能做，真正意义上的投入成本几乎为零。

中国市场那么大，企业那么多，需要 SEO 服务的公司和网站也数量巨大。当然这些公司还不一定意识到自己需要 SEO 服务，需要你去说服他。

刚入门的 SEO 服务商也不需要很有名。可以先从自己身边寻找可能的客户，亲戚朋友、同学、本地企业，先免费或低价做一两个客户，客户又可以向其他人推荐，要找到一些客户并不很难。

成为一个 SEO 服务商确实是一件很简单的事情。这也就是现在网上 SEO 服务泛滥，水平层次不齐，价格天差地别的原因。对整个行业来说，这不一定是件好事，必然造成良莠不齐，SEO 服务的整体形象也可能被连累。但是对一个想要迈入 SEO 行当的人来说，这就变成了优势。

2．SEO 服务的劣势

从另外一个角度说，提供 SEO 服务有天生的致命缺陷，其中最主要的是，SEO 服务难以扩展。SEO 服务基本上是人工去做，虽然网站流量分析、排名跟踪、外部链接的查询跟踪，甚至网站基本优化水平的判断，都有很多软件可以协助自动化，但是针对一个特定网站，寻找到需要优化的地方，提出方案，执行优化，创造性的建设外部链接，这些都只能人工完成。

每个网站有自己的特点，不同行业有不同的竞争情况和要求，所以网站的优化方案全都不一样，没有一个适用于所有网站的模式。再加上 SEO 不确定成分很大，谁也无法确保排名一定进第一页，也无法保证获得的排名不会下降。SEO 过程中需要很多说不清道不明的经验、直觉，这就更无法自动化。

所以说到底，提供 SEO 服务是在出卖自己的时间。凡是出卖时间的生意，都注定扩展性很差。如果服务一个客户需要一个人，服务两个客户就需要两个人。这与生产和销售大部分实物产品很不一样。生产一百件衣服可能需要一个人一天时间，但生产一千件可能只是需要两个人一天时间，而不会是十个人。

这种低扩展性决定了 SEO 服务有内在风险。业务好时，服务商不得不扩充公司规模，没有其他方式解决。一旦业务量下降，人员成本可能就成为公司的巨大负担。就算没有扩充计划的个人 SEO 服务提供者也一样受限制。一个人的时间就是那么多，注定不可能同时服务很多客户。

这里给有志于 SEO 服务的站长提几条建议。

（1）清楚定位

自己擅长什么？目标是高端市场还是低端市场？是专做关键词排名还是整站优化，专注于流量？这些都必须有一个清醒的认识，才能使自己的服务有特色。

（2）考虑 SEO 细分市场

SEO 说起来简单，但其中每一个过程要想做精通也不容易，所以可以专注于特定部分，做专做精才有号召力、吸引力。比如：

- 专门做关键词研究
- 专门做链接诱饵
- 专门做 Follow 链接的论坛、博客链接发布
- 专门做某个行业（如医药、旅游行业）的 SEO
- 专门做某个 CMS 系统或购物车的 SEO
- 专门做 SEO 软件等

不要小看了这些细分市场，每一个细分市场都有巨大容量。

（3）卖出一个好价格

前面说过，SEO 没有良好扩展性，也就没办法以数量取胜。要想提高利润，就只能提高价格。实际上像 SEO 这种相当专门化的服务，是很有可能卖出比较高的专业价格的。

网上很多所谓 SEO 提供商以几千元甚至几百元的价格优化一个网站，这确实是太低估了自己的价值。大家可以对比一下，一个好的补习老师每小时收费都可以达到几百元，SEO 是更少人懂得和精通的技术，一天的服务时间定价几千元人民币一点都不过分。从客户角度出发，如果 SEO 服务能让一个商业网站流量翻倍，那么给网站带来的销售和利润是多少？很可能是几十万元、几百万元甚至上千万元。这样的 SEO 服务却收费几千元就太不正常了。

（4）流程化

虽然大部分 SEO 工作无法自动化，但是有一定规模的服务商应该尽量做到流程化。公司内部完整配备 SEO 人员，清楚划分工作范围，项目必须制定任务清单、时间表及考核标准。随着越来越多的网站认识到 SEO 的重要性，在搜索引擎上的竞争也越来越白热化。靠感觉、靠一个人的力量已经很难适应 SEO 服务的市场，团队作业是一个最好的方向，而流程化是团队作业不可缺少的要素。

10.3.2 给自己做 SEO

SEO 的第二个、也是更光明的出路是给自己做 SEO。我在 2006 年开始写 SEO 每天

一贴时就说过，SEO 的最终出路是做自己的电子商务网站。现在这个观点已经被很多 SEO 人员接受和认可，并跟我说几年前我就提出这样的观点很超前。其实这个观点一点也不超前。SEO 从诞生那天开始，就主要是为了给自己网站做优化、带来流量的。研究 SEO 的前辈们都是从解决自己网站排名和流量问题开始的，哪会首先想到给别人做 SEO。

SEO 只是一个工具，并不是目标。但 SEO 是最强大的网络营销工具之一。掌握 SEO 就掌握了网站赚钱最重要的要素之一 —— 流量。给别人做 SEO，自己只能获得 SEO 价值的一部分，常常是一小部分。如果你能帮其他网站流量翻倍，利润翻倍，你作为服务提供商，服务费只能是利润的十分之一，甚至百分之一。如果你是给自己做 SEO，你就能拿到利润的大部分。

自己做网站，既可以做信息类网站然后卖广告，或通过联署计划赚取佣金，也可以直接做电子商务网站卖产品。现在国内网上购物已经非常普及，支付、快递两个大问题已经基本解决，做自己的 B2C 网站是大势所趋。

如果英文好，做外贸网站也是一个非常好的方向。中国是制造大国，各种物美价廉的产品应有尽有。掌握货源是国内公司和站长的巨大优势，再加上通过 SEO 带来流量，你将如虎添翼。

很多 SEO 人员对 SEO 行业的前途感到迷茫，其实退一步把 SEO 作为一个工具，你就会豁然开朗。传统 SEO 服务提供商现在也纷纷开始做自己的电子商务网站，既能赚取更多利润，无限扩展业务范围，也能降低风险。做自己的网站甚至对 SEO 服务反过来也有促进作用，比如对于太过挑剔，把价格压得很低的客户，你就可以很有底气地拒绝。与其花时间做收费低廉的 SEO 服务，还不如做自己的网站，这样你才能避免成为廉价劳工。

10.4 SEO 不是免费的

很多人在讨论 SEO 的优势时会提到，SEO 是免费的。其实 SEO 并不免费，也是要付出成本的。这里所说的成本不是指使用其他公司来优化网站的服务费用。就算你的网站不用任何外部服务，也不买任何软件，完全自己动手优化，也还是要付出成本的。有时这个成本还很高。

10.4.1 人力成本

最显而易见的是人力成本。网站本身一般来说只要大规模优化一次，以后再小幅度修改即可，这部分的人力成本也许可以算入技术部门。但是外部链接建设、网站流量的跟踪、SEO 策略更正、发现热点建设专题内容、遇到问题时分析寻找原因等这些工作，都不是一次性的，需要常年进行，这就必然需要有专门的 SEO 人员。对一些大网站来说，可能还需要一个 SEO 团队，工资等人力成本往往比外部服务费还要高很多。

10.4.2 机会成本

另一个不容易直接看到的是时间及机会成本。通过 SEO 做流量是需要一段时间的，尤其是新网站。不要指望几个月内有很好的流量，对 SEO 的预期效果至少要放在半年到一年之后。但网上机会稍纵即逝，某些当红类型的网站要想迅速占领市场，靠 SEO 推广是不现实的。等过了半年一年，早就过这个村没这个店了。想象一下推广诸如凡客诚品、开心网这类网站，SEO 根本就不是一个选项。等你通过 SEO 做出一定流量，竞争对手早就占领市场和用户的头脑了。

所以 SEO 只适用于那些耗得起时间的网站，半年到一年之后，市场不至于出现惊人的变化。

10.4.3 失败风险

最终是否有效果也是 SEO 的一个风险。搜索引擎不是我们开的，谁也无法百分之百确保做了 SEO，就一定能有排名和流量。再厉害的 SEO 也不能百分之百肯定。相信一些网站有这样的经验，花了时间、人力、精力、服务费，却没有什么效果。

这与 PPC 是不同的。PPC 只要花了钱就一定有流量，SEO 做了却没有效果是经常发生的。在这种情况下，使用 PPC、网络广告、事件营销、口碑传播等要比 SEO 划算得多。

10.4.4 SEO 成功风险

就算通过 SEO 得到了很好的流量，也可能导致过度依靠搜索引擎流量的风险。我看过一些网站，搜索流量占到总流量的 80%～90%甚至更高。这一方面说明 SEO 做得不错，另一方面也意味着很大的潜在威胁。一旦搜索引擎改变算法，哪怕并不是针对你的网站有任何惩罚，也可能造成流量急剧下降甚至消失。如果没有点击流量和直接访问流量作为平衡，很可能对网站是一个致命性的打击。SEO 做得好的公司和站长，还必须花费更多精力去开拓流量来源，千万不可过度依赖搜索流量。

10.5 不要做奇怪的事

经常在论坛看到或者有朋友直接问我一些奇怪的问题，有的是比较大的涉及整个网站的，有的是很小的代码细节问题。看到这样的问题，我的第一个反应经常是好奇：为什么要做这种奇怪的事情呢？

比如有的站长会把同一个域名的网站内容改成完全不同的行业或话题，诸如从减肥信息网站转眼之间改成股票投资知识。有些域名确实没什么内容针对性，比如数字组成的域名，但是这不意味着整站改变主题会有什么好结果。进一步问对方为什么要这么大

幅度地改变内容，原因常常是一些无关紧要的小事，诸如旧内容做了几个星期没看到效果。可以肯定，抱着这样的想法换成新内容，过几个星期还是不会看到效果。

搜索引擎面对这样奇怪的网站也会犯糊涂，这个网站究竟是讲什么的呢？抓不住网站的主题，自然无法给予好的排名。我们经常看到这样的现象，一个网站的内容更改以后很长时间，还会在搜索原来的关键词时看到排名。这说明搜索引擎对网站历史数据有记录，并不会因为网站当前页面内容彻底改变而完全忘记以前的内容。

再比如，还有人问过，在 http://www.abc.com 和 http://abc.com 放上不同的内容会怎么样？技术上来说，www.abc.com 是 abc.com 的子域名，http://www.abc.com 和 http://abc.com 这两个 URL 完全可以放不同的内容。但是无论搜索引擎还是用户都约定俗成，这两个 URL 应该是同一个网站。搜索引擎访问这两个 URL 时，如果都能正常返回 200 代码，还会自动选一个 URL 版本作为网站的规范化网址。这也就是通常建议选择一个版本的原因所在。这两个网址放上不同的内容，搜索引擎又会被搞糊涂。

再比如一些细节问题。有的 SEO 喜欢把图片放在 H1 标签中，认为这样会增强图片 ALT 文字的权重。实际上这也是一件比较奇怪的事。把图片放在 H1 标签中，用户在浏览器上看不出任何区别。仅从代码上看，你也可以把图片放在黑体标签中，用户也看不出任何区别，图片又不能加粗。所以这样的代码对用户毫无意义，搜索引擎也会琢磨这样做意义何在。

诸如此类奇怪的做法，只能把搜索引擎搞糊涂。

网站之所以需要 SEO 的其中一个原因就是，让搜索引擎能更快速、更容易地提取网站内容，千万不要给搜索引擎设置障碍，让搜索引擎自己琢磨是怎么回事。SEO 要做的是尽量减少搜索引擎的工作量，让网站内容在搜索引擎眼里直接明了、符合常规。在网站上做些奇奇怪怪的事，搜索引擎要么被搞糊涂，要么也以奇奇怪怪的方式处理。

10.6 解决基本问题就解决了 95%的问题

网上 SEO 信息汗牛充栋，读者您现在又在看这本厚厚的 SEO 方面的教程，涉及的 SEO 问题方方面面，显得颇为繁杂。在实践中，几乎没有网站能面面俱到，优化本书中提到的所有方面。好在 SEO 人员也完全没有必要解决百分之百的问题。一般来说，解决了最基本的问题，就解决了 95%的 SEO 问题。

这里所说的基本问题主要包括：

- 至少整体方向的关键词研究。
- 网站结构及内部链接，解决收录问题。
- 页面标题标签、H1 标签等几处最重要代码。
- 页面正文包含关键词两三次。

- 网站内容原创性和避免复制内容。
- 找到一两个最拿手的外部链接建设方法。

其他诸如关键词密度、关键词出现的位置、链接锚文字、URL 的设计等，花很少时间就能优化完成当然最好。如果非基本问题需要花很多时间、精力才能解决，也大可暂时搁置。把时间花在刀刃上，往往解决了少数基本问题就能看到 95%的效果，再花更多时间解决细枝末节的技术问题，对排名和流量很可能只能贡献不超过 5%。与其花大把时间在小问题上面，还不如去建立新的网站，或者专心扩展内容、建立外部链接。在这个意义上，SEO 其实是很简单。

10.7 自然和平衡的艺术

我一直觉得 SEO 更多地是一门艺术，而不是技术。

相信做网站设计和 SEO 的以学理工科的居多，但是真正的 SEO 所要求的文科基础更多，比如市场营销、广告、心理学、写作等。当然，它也要求一些技术基础，比如 HTML、编程、服务器基础知识等。但对这些技术内容的要求不是很高。一个高级程序编写员并不一定能成为一个好的 SEO 人员，甚至一个搜索引擎工程师也不一定能成为最好的 SEO 人员。

说 SEO 是自然的艺术，指的是对网站的优化应该看起来是自然而然的，无论搜索引擎还是用户都不应该感觉到你对网站做了优化。

说 SEO 是平衡的艺术，指的是不能把 SEO 手段用到极致。对一个网站必须有总体的考量和计划，不同的推广渠道，不同的 SEO 手法，要取得一个均衡，不应该有任何一个因素显得非常突出。

当然，这是最理想的情况。实际上这两方面都是说起来容易，做起来困难。SEO 专题一章讨论过避免过度优化，但怎样才算过度是很难判断的，既没有公式可套，也没有明显特征可观察。有很多人觉得 SEO 很简单，读过一些资料后就成专家了。实际上一些细微的地方需要积累很多经验，对搜索引擎和 SEO 有深刻的理解，才能找到感觉。

自然和平衡可以体现在 SEO 的方方面面。举一个自然性的例子。我们都知道登录分类目录是获得外部链接的一个常用方法。有的站长会在网站设计完成后，集中时间向所有能找到的分类目录提交，这也是很多专家所推荐的。但在实际工作中，我建议不要同时向这些目录提交。

站在搜索引擎的角度看，如果某一个网站突然在某一天或某一段很短时间内，在所有重要的网站分类目录中出现，你觉得这自然吗？很显然，只有脑子里想着 SEO 的人才会这么做。一般的生意人会想到大规模向分类目录提交吗？绝大部分不会。

其他外部链接建设方法也是如此，必须循序渐进，持之以恒，让外部链接保持匀速自然增长。不要专门拿几天时间搞突击，增加大量外链，然后一段时间又不再做外链建设。

再举一个平衡性的例子。一般来说，最好的优化是使网站的很多关键词排名提高，而不是只有几个最关注的关键词排到第一，但其他相关关键词和其他网页都排得很靠后。也就是说，不应该有一个全明星网页（一般来说是主页）在搜索行业最热门关键词时排第一，而其他相关关键词都找不到你的网站。

这也同样适用于权重在整个网站的均衡分配。通过适当的内部链接结构，将主页 PR 和权重均匀地分散到各个网页，使所有网页的权重得到提升，既有利于收录，也有利于排名能力。就算有几个网页是你想侧重的，也不应该让这几个重要网页和其他网页的悬殊太大。

10.8 SEO 是长期策略

缺乏耐心是 SEO 的大忌，却是在 SEO 新手身上最常看到的心态。

经常在 SEO 论坛看帖子的人都会碰到一些站长问，自己的网站没有排名、没有流量很着急，该怎么办？再深入问一下对方网站运营多久了，才发现域名刚注册两个星期，站长就已经心急如焚了。对这样的站长，除了劝他多等几个月，也没有什么好建议了。

10.8.1 实施 SEO 需要时间

任何事情都有自身的规律，不以人的意志为转移。SEO 是一项长期策略，除了一些特例，比如竞争程度比较低的关键词，或大公司网站、引起社会关注的网站，很少能快速看到效果。这是搜索引擎本身原理造成的，与 SEO 人员自身关系不大。

对有一定规模的网站而言尤其如此。我们姑且不论优化整站需要花的时间、精力，从完成优化算起，搜索引擎把一个大中型网站页面重新抓取一遍，往往就需要几个月的时间，再进行索引，重新计算权重分配，还要加上外部链接的收录，计算域名信任度的累计，这些都需要时间，而且是不短的时间。所以正常的 SEO 策略都是以至少半年到一年为周期的。

一些公司网站并不是 SEO 人员没有耐心，而是老板不允许有耐心，老板们希望立即看到效果。SEO 也应该实话实说，SEO 就是需要比较长的时间，可以说这是它的劣势。要想立即看到效果，请做 PPC。对个人站长来说，没有上层的压力，完全可以耐心地等待。实际上我经常有这样的体会，一个网站做好后放在那里不再管它，既没有内容更新，也没有链接建设，过两三年后排名和流量自然而然就上升了。

10.8.2 不进则退

网站排名和流量提高并稳定以后，也不能说大功告成。SEO 是个没有结束的过程。有一些网站不再继续做 SEO，排名可以稳定很长时间。有的网站却不进则退，停止优化就很可能导致排名和流量慢慢下降。

造成排名下降的原因有很多方面，首先是搜索引擎算法的不断改进。Google 每年都有几百次小的算法变化，平均每天不止一次。百度也几乎每星期都有变化。以前有效果的 SEO 技术，现在可能不再有效。以前大家都不注意的地方，搜索引擎可能悄悄增加其在算法中的权重。比如四五年前外部链接数量最重要，后来慢慢转化到以质量取胜，再后来锚文字越来越重要，近期锚文字权重似乎又有所下降。页面上的优化同样如此，以前 H1 文字权重比较高，近两年有所下降，图片 ALT 文字权重有所上升。这些细小变化都会引起网站排名、流量的波动。SEO 人员必须长期积累，关注搜索引擎算法改变，必要时对网站做出改进。

其次，竞争对手也在不断提高。任何有商业价值的领域都会有多个竞争对手。你的网站排到前面就停止优化，竞争对手却不会闲着。他们在研究你的网站，不断修改，增加外部链接。不进则退，SEO 人员必须始终关注竞争对手的动向，保持领先。

用户的搜索习惯也常常随时间而改变。几年前搜索用户还没有经验，喜欢搜索比较短的词，现在大部分网民都知道搜索长的词组甚至句子能得到更准确的结果，平均搜索词长度一直在稳步增长中。

某些行业受关注度也有起有落。比如搜索“SEO”这个词的人数近几年就在稳步上升中。同样与网站相关的“网站设计”之类的词，则没有提高，甚至还在下降中。SEO 人员必须关注用户习惯和社会焦点的变化，思考网站是否需要增添新的内容，抓到新的社会热点。

整个互联网每时每刻都在变化。每天都有新网站出来，也有网站关闭，甚至完全消失。以前做的外部链接可能因为对方网站消失，或者页面被推到更深层的内页，以至于不再被收录或权重下降。

整个网上的变化 SEO 人员无法控制，只能从自己网站出发，不断做大做强，才能立于不败之地。

10.9 没有 SEO 秘籍

看完本书，读者大概会感觉到 SEO 是一件挺辛苦的工作。要做好 SEO，提高网站流量，没有捷径，只有踏踏实实做好基本优化，扩充内容，吸引链接。

10.9.1 为什么没有 SEO 秘籍

有的 SEO 初学者以为有什么 SEO 秘籍。经常在论坛看到有会员问一些问题，没有人给予明确回答时，大家会怀疑高手们有所保留，不愿意透露自己的秘密。其实大家尽可以放心，根本没有 SEO 秘籍。高手没回答你的问题，不是想保密，而常常是问题太宽泛无从回答。

我接触的国内外 SEO 人员应该算比较多的，个人也接触了不少大大小小的 SEO 项目。据我所知，没有所谓 SEO 秘籍。如果你在网上看到谁在出售 SEO 秘籍，或者提供绝对百战百胜的 SEO 培训，你购买或去参加的话，注定会失望而归。如果真的有 SEO 秘籍，不用说几百元，几百万元也不会有人卖给你。如果真有人掌握了搜索引擎算法秘密，想做什么关键词就做什么关键词，把搜索引擎玩弄于股掌之间，那想赚几百万几千万都是轻而易举的，谁还有功夫卖几百几千元的秘籍呢？

10.9.2 搜索引擎排名算法的秘密

实际上掌握搜索引擎全部核心算法的一共也没有几个人。就算是搜索引擎工程师，也只是了解自己负责的那一小部分。世界上知道搜索引擎全部排名秘密的，大概不会超过一二十个人，不然每年从搜索引擎离职的工程师人数也不会太少，这些工程师如果都知道排名秘密，这些秘密也早就泄露出去了，或者早就出现了几个百战百胜，玩弄搜索引擎于股掌之间的 SEO 公司或电子商务公司。实际上并不存在这样的公司，就连最顶级的 SEO 团队和商务网站也经常搞不懂排名变化的原因，也在不断研究和尝试中。

Matt Cutts 是 Google 最资深的工程师之一，他在博客和评论中也经常提到，具体算法问题得询问其他同事。有的时候 Matt Cutts 发了帖子后发现并不准确，还要更正。连这个层级的工程师都有很多不知道的算法细节，就更不要说做 SEO 的了。

现在的搜索引擎算法有这样一种趋势，那就是基本规则相对透明化。但是就算 SEO 和所有站长都明知道这些算法和规则，也很难用来作弊和人为控制。比如说我们都知道高质量链接非常有效果，但是得到高质量链接却是一件很困难的事，不是靠作弊能做出来的。所以搜索引擎算法中能被利用的漏洞越来越少。

10.9.3 SEO 绝招

SEO 没有秘籍，不过可以有绝招。这种绝招是建立在对大家都知道的基本规则的深入理解和经验上的。比如大中型网站的收录问题是 SEO 的一个难题。其基本原理非常简单，可以归纳为两条：一是页面必须有导入链接，搜索引擎才能发现页面，才能被收录；二是距离网站首页点击距离越短越好。对一个中等权重的网站来说，点击四五次之内的页面应该可以被收录，点击七八次以上才能达到的页面就很难被收录了。

了解基本原理，却不意味着 SEO 人员能解决收录问题。面对一个有一定规模的网站，栏目或分类众多，随时有新内容加入，要怎样调整网站内部结构和链接关系，使所有页面都符合上面两条要求，各有各的绝招。

再比如建设外部链接。有的人纯熟掌握一两种绝招，诸如链接诱饵或 CMS 系统模板，把它用到极致，在外部链接上就能战胜绝大多数竞争对手。

这些绝招都不是建立在秘密之上的，而是建立在大家都知道的规则上的。

10.10 SEO 不仅是排名

对 SEO 缺乏了解的新手最常有的误区之一是把 SEO 等同于关键词排名。其实关键词排名仅仅是 SEO 的一部分，而且是比较初级的部分。真正的、全面的 SEO 所包含的内容比关键词排名要广泛的多。这一点大部分稍有经验的 SEO 都知道。不过口头上说是一回事，真正优化网站时很多人又不自觉地把关键词排名作为目标。

做过大中型网站的人一定都会有这样的感触，真正带来大量流量的是长尾关键词，而不是自己设想的那几个主要目标关键词。我所接触的 SEO 项目几乎都经历这样一个过程，网站整体优化之后一段时间，网站流量有质的飞跃，但是最主要的几个关键词排名没有明显提升。

本书案例部分的网站就是一个很好的例证。从 2009 年 10 月份开始逐步优化，2010 年 1 月 1 号上线第一个比较完整优化的版本，上线后又发现了不少没有预见到的问题需要改正，到 2010 年 2 月份时，优化还没有全部完成，不过 2010 年 2 月份搜索流量已经增长了二十倍以上。但是我个人监控的 50 多个关键词排名绝大部分没有明显变化。带来流量的是那些靠人工不可能预期到的长尾关键词，既包括由于网站结构优化而收录的新页面，也包括页面优化造成的更多长尾关键词得到排名。

当然，提倡整站优化并不意味着不能去做关键词排名。如果搜索量大的关键词排名好，也可以带来很多流量及品牌效应。不过执着于特定关键词排名是一个不确定性高、难以规模化的方法。通常建议的方法是先做整站优化，带动长尾流量，主要关键词慢慢去做。时间久了，整站实力提升，自然也会带动主关键词。

说 SEO 不仅是排名的另一个原因是，由于地理定位及个性化搜索的原因，不同的人搜索相同关键词看到的结果并不相同，尤其是地理定位因素现在已经很明显。个性化搜索目前影响还比较小，但今后几年有可能日趋明显。这就从根本上瓦解了关键词排名的意义，因为已经不存在一个大家都能看到的相同排名。

另外，有了排名，也不一定意味着有点击和流量，更不意味着一定能带来订单。排在前面的页面要被用户点击，还要受品牌效应、页面标题吸引力等因素的影响。

如果遇到潜在客户问：某某关键词排到第一页收多少钱？遇到这种咨询，首先要做的不是按照他们的要求报价，而是先扭转对方的观念。

10.11 SEO 不是作弊

近两年，SEO 越来越受到互联网行业人士的关注和重视，但毋庸讳言的是，SEO 在很多网民心目中名声并不好。在很多人眼里，SEO 就等同于作弊，一提起 SEO，大家就想到群发垃圾信息和关键词堆积。网上一有人提到博客上的垃圾留言就会有人说，又是那些 SEO 人员干的。就连很多互联网从业人员也持有这种想法。这应该是早期 SEO 作弊、垃圾泛滥给大家造成的负面印象，也算是咎由自取。

其实 SEO 不是作弊。

网上有一种说法，认为 SEO 是在网站上做手脚，钻搜索引擎的空子，是把一个只能得 50 分的网站，装得看起来像是 90 分，从而提高排名。真正的 SEO 并不是如此的。SEO 的目的是做强做大网站，不是把 50 分的网站伪装成看似 90 分，而是把 50 分的网站补强为确实能得 90 分。

读者们从本书就可以看到，SEO 涉及的内容很广泛，绝不是在页面上堆积一些关键词那么简单。做关键词研究和竞争对手调查，也就是在了解用户需求，扩展网站内容也就是满足用户需求。良好的网站架构和内部链接是提高用户体验很重要的方面，页面的优化也使用户更容易抓住内容重点、易于浏览。流量分析和策略改进，相当于重视用户反馈意见，改进网站。外部链接建设也是另一种形式的公共关系，与其他网站和用户更有效地互动。所以说，SEO 的整个流程无不是为了把网站做好做强，与作弊有天壤之别。

还有的人认为 SEO 违反搜索引擎规则，其实恰恰相反，SEO 是搜索引擎的朋友。搜索引擎要给用户提供有用的信息，就必然需要信息源。搜索引擎友好、易于收录、主题突出、内容丰富的网站是为搜索引擎提供信息源，满足用户需求，搜索引擎求之不得。如果网上的网站净是些不容易收录的页面，搜索引擎才真正头大了。

也正因如此，Google 有很多员工在博客、论坛中很活跃，发布信息、回答问题，积极参与搜索引擎营销行业大会，以某种形式指导站长做 SEO。百度也于 2010 年 4 月份开通了半官方的百度站长俱乐部，针对 SEO 有关问题做出回答。Google 和必应都发布过 SEO 教程。主要搜索引擎从来没有反对 SEO，甚至可以说支持白帽 SEO。

而且搜索引擎本身也在做 SEO。雅虎、MSN 都公开招聘 SEO 人员，优化他们自己的网站。

要扭转 SEO 在普通网民心目中的形象，还要靠 SEO 从业人员自己的努力，不要作弊，坚持白帽，做好网站，提升用户体验。

10.12 SEO 与网站运营

SEO 是网络营销利器。SEO 做得好不好，有时候是一个网站能否成功的关键。

但 SEO 只是网站运营的一小部分。SEO 人员不仅要了解 SEO 技术，也要梳理好 SEO 在整个网站运营中的地位，不能喧宾夺主。影响网站运营的因素太多，SEO 只是其中很小的一部分，很多时候是非必需的部分。

SEO 只是搜索营销的一种，搜索营销还有 PPC 形式。搜索营销又只是网络营销的一种。网站完全可以抛弃搜索营销，抛弃 SEO，靠其他网络营销方式带来高质量流量，比如电子邮件营销、联署计划营销、论坛营销、口碑和病毒式营销、博客营销、社会媒体营销、电子书营销等。任何一个网络营销方法使用得当，都可以成就网站流量。SEO 虽然是最有效的流量建设方法之一，但绝不是唯一的方法，并不是每个网站都要做 SEO。

网络营销也只是网站运营的一部分。除了流量，运营网站还要处理很多其他事情。产品的研发或采购、物流管理、人员管理、资金流动、品牌建设、网站技术系统研发等，任何一个环节做不好，都有可能导致网站失败。

所以，有经验的 SEO 一直强调，给自己做 SEO，卖自己的产品是最好的出路，但实施起来难度也不小。SEO 做好了只是解决了流量，却不能解决产品、物流、客服、抗风险等一系列问题。很多时候单纯从 SEO 出发，会与网站运营的总体目标相冲突，这时 SEO 就要服从网站运营总体规划，毕竟，光有流量是没有用的。

10.13 内容为王

这里所说的内容指的是高质量的原创内容，而不是转载甚至抄袭的内容。

10.13.1 原创内容是 SEO 的根本

内容为王，链接为后，这个说法 SEO 人员应该都很熟悉。其原因不仅仅在于搜索引擎给予原创内容和外部链接很高的排名权重，更在于这两者是 SEO 最大的难点。网站结构、内部链接、页面优化、关键词分析、流量分析，这些重要的 SEO 步骤大体上都在 SEO 人员控制范围之内，一个有经验的 SEO 可以相对顺利地完成。但是高质量的内容和外部链接，往往超出 SEO 的控制，经常是可遇而不可求的。

互联网的发展使我们进入真正的信息爆炸时代。不过仔细审视一下又可以感受到，这也是一个信息匮乏的时代。网上转载、抄袭、复制的低质量内容太多太多了，真正有价值的原创内容却太少。

中小网站站长都会有这种体会，面对一个或一系列产品，总感觉没什么内容可写。

产品说明写来写去都差不多，找不到原创的切入点。对一个有决心有毅力的 SEO 来说，这也是机会。别人都不写，你写了就占据了很大的优势。

高质量的内容首先符合搜索引擎的目标。前面说过，搜索引擎的目标是给搜索用户提供信息，而不是给你的网站带来客户。只有你的网站提供了内容，才能与搜索引擎的目标达成一致。

原创内容也是页面收录的重要推动力。不少站长都为网站收录问题发愁，搜索引擎蜘蛛要么只访问首页，不再进一步抓取内页，要么收录之后又被删除。缺少原创内容是关键原因之一。一个以转载、抄袭为主体内容的网站对搜索引擎来说毫无价值。收录这样的页面越多，越浪费搜索引擎的带宽、数据库容量、爬行时间、计算时间，却不能给搜索用户提供更多信息，所以搜索引擎根本没有收录这些内容的理由。

原创内容也是外部链接建设的重要方法。除非是交换链接，不然其他站长没有理由链接到一个没有实质内容的网站。

10.13.2 内容策划是 SEO 策略

内容为王的道理大家都清楚，但在实践中往往又很迷茫。花时间、精力原创内容到底值不值得？这一点我们从一个缺乏内容的网站能看得更清楚。做中小型电子商务网站的人都会感受到做 SEO 非常困难，原因之一就是中小电子商务网站基本上没有独特的内容。就算网站有成千上万的产品，但整个网站还是内容匮乏。除了首页可以想办法写几段原创内容外，分类页面都是产品名称加链接，产品页面无非就是产品说明、型号、图片，看似有不错的内容，但其实大部分产品说明都是从生产商那里照搬过来的，网上已经有成百上千的网站在使用相同的产品说明。如果站长仅仅把来自生产商的内容放上去，很难看到 SEO 效果。

在这一点上，中小电子商务网站无法与有实力的大型电子商务网站相比。大公司网站就算没有实质内容，也可以靠公关、广告把域名权重做得很强。这种高权重域名上的复制内容，反倒会被当成原出处得到好的排名。

要解决这个问题，不是修改代码这种技术性的优化能做到的，而必须上升到 SEO 策略层次，也就是说在策划网站时就要计划好怎样产生原创内容。有几种 SEO 内容策略可以考虑。

比如中小电子商务网站分类页面，必须人工添加分类说明文字，长度至少在两三段文字以上。虽然要花时间，却是不得不做的。好在分类页面应该不会太多。

再比如在电子商务网站上增加用户评论功能。看看亚马逊书店的页面就知道，用户评论是非常重要的内容策略。调动用户积极性，用户贡献的内容既是绝对的原创，又能无限扩展。

还可以增加行业信息、商品导购、员工博客等部分，在重复的产品说明之外，增加原创内容。

有的网站在设计功能时就以用户产生内容为主，如分类广告、B2B、论坛、社交网站等。这类网站只要设计好搜索引擎友好的平台架构，用户就自动为网站创建了内容。

SEO 人员还可以花时间在优化用户发布内容的流程上，指导用户怎样写最适合的标题、正文及标签，选择哪个产品分类最合适。

所以内容设计在网站一开始就要进行，对缺少原创、也缺少用户产生内容的电子商务网站来说，更要想办法增加原创内容。

10.13.3 内容推广

另外一个需要注意的是，有了好的内容并不意味着别人就会自动知道，要让用户发现内容，让其他站长看到有意思的内容进而链接过来，你需要做初始推广。仅有好的内容，缺少最初的推动，也很难发挥原创内容的威力。不过初始推广其实是很简单的，关键在于找到行业中几个有话语权的人，让他们知道你的网站。这些人一旦发现你确实有好的内容，一定会影响到更多人看你的网站。

很多人对这一点有所怀疑。据我的经验，这种方法屡试不爽。觉得这种方法没有用的人其实都没有真正认真去尝试。

如果你的行业没有几个权威人物存在，你也可以到相关的论坛上深入参与讨论一段时间，在不唐突的情况下稍微推广一下你的网站，就会有人注意到。越是别人都不愿意做原创内容，大家都看不到好的内容，你就越有机会通过简单的初始推广，发挥原创内容的效力。

10.14 具体问题具体分析

和其他任何事物都一样，SEO 也需要具体问题具体分析，不存在放之四海而皆准的公式。本书中介绍的 SEO 技术，或者在网上、其他书里看到的技术，都只能是就一般而论的，告诉你 SEO 的规律性技术。但每个网站都有每个网站的具体情况，面对一个特定网站，该怎样优化？既不能完全照书本，也不能盲目套用其他网站的手法。

影响网站具体技术运用的因素如下。

- 网站本身情况：是大网站还是小网站？内容是原创为主？还是只能转载为主？CMS 系统是开源的？还是自己设计的？产品分类有多少种？能否使网站尽量扁平化？还是产品数量巨大，怎么做也不可能扁平化？
- 公司情况：公司是否有足够的决心做 SEO？能投入多少人力物力？能耐心等待多

长时间？技术部门是否愿意配合？内容编辑部门是否具备基本 SEO 知识？公司公关部门实力如何？SEO 是否得到公司高层的足够重视？

- 市场及竞争对手情况：与最强的竞争对手外链、排名差距到底有多大？只有一两个有实力的竞争对手，还是有一二十个？整个目标市场容量有多大？是否有足够的关键词搜索次数？

这些因素都将决定一个网站到底该采取什么样的 SEO 策略和手段。

举个例子，在关键词研究部分说过，确定关键词的宗旨是搜索次数多，竞争者少，相比之下最容易使 SEO 效能最大化的词是最合适的关键词。但在实际操作中，这并不是唯一的选择。如果公司预算够多，公关部门实力强，又愿意配合 SEO，公司高层决心大，那么也不妨把关键词定位在搜索次数最多的词上，而不管竞争对手情况，不达目的誓不罢休。如果你是个人站长，也可以采取这种策略，因为你有足够的时间和耐心，可以等上几年。而对一些能投入的资源少，却又必须尽快看到结果的公司来说，就只能按一般规律找到效能最高的关键词。

关键词在网站上的分布也同样需要具体问题具体分析。关键词研究章节讨论过，不同热门程度的关键词应该分好层次，与网站的首页、目录页相对应，均匀分布在整个网站上。但有的时候就没办法做到关键词的完整规划。比如案例部分的网站是一个比较购物网站，页面包罗万象，不针对特定产品或市场。而分类又无法按关键词热门程度而定，因为整个网站包括了几乎所有网上能卖的产品，每个大类都有特别强的对手，无论是首页还是一级分类页面，都可以肯定无法与相应分类的市场领跑者竞争。所以产品分类页面的关键词几乎可以说只能暂时放弃，必须把精力放在产品页面和长尾关键词上。

优化网站的其他步骤同样如此，必须根据网站的特定情况，选择最适合的方法。有的时候，书上教你不要用的方法，可能正是最适合你的方法。

第 11 章
SEO 工具

做网站优化的一个瓶颈是 SEO 的自动化。到目前为止，绝大部分网站优化工作还得是人工去做。一些 SEO 工具可以辅助，但还没办法取代人工操作。

SEO 工具软件大致可以分成四类。严格来说只有前两类可以算是真正的 SEO 工具，不过后两类也被讨论得很多，姑且提一下。

1．SEO 信息查询工具

包括线上工具和可以下载、运行于客户端的软件，主要是查询一些与 SEO 有关的数据，包括排名位置和网站基本信息，比如 PR 值、关键词密度、关键词排名、反向链接数等。

这种查询工具对 SEO 的前期调查及效果监控是必不可少的，对提高工作效率无疑很有帮助，而且准确性很高，与自己手动查询没什么大区别，又能省很多时间。

搜索引擎并不喜欢大量工具自动查询，这对他们的资源是个浪费。不过只要别太过分，限制一段时间内的查询次数，一般问题不大。如果来自一个 IP 的自动查询过多，搜索引擎可能会暂时屏蔽这个 IP 地址。

本章介绍的 SEO 工具主要是信息查询工具。

2．网站诊断工具

这类工具很少见，制作起来不容易，也很难准确。由于搜索引擎排名算法的复杂性和变动性，诊断软件给出的建议最多只能作为参考。

最著名的诊断软件是 Webposition Gold，会自动分析用户网页与排在前面的网页之间的差别，然后给出优化建议。Google 点名道姓地指出不欢迎使用 Webposition Gold。

网站诊断工具还很不成熟。比如软件抓取目标网页，进行分析之后可能会告诉站长，需要把关键词密度提高到多少多少，标题中关键词重复两次或者三次。这些建议无非是对相关关键词排名前十位或二十位的网站进行统计得出的。

问题在于这些统计数字其实是有误导性的，它欠缺了一个好的 SEO 人员应有的全面观察、直觉和经验。我们看到的排名与页面元素之间并不具备因果关系，虽然可能呈现某种统计相关特征。

举个例子，查询某关键词，排在前十位的网页标题平均出现关键词两次，而排在比较靠后的网页标题平均出现关键词次数比两次高或者低。这是不是意味着在标题中重复两次关键词是最优化的呢？乍看之下是这么回事，但仔细思考一下，这两者之间并没有因果关系，前 10 名的网页很可能是因为其他因素排在那里的。更何况排名往往和页面特征没有明显关联。

所以，目前的网站诊断软件给出的某些建议可以采纳，比如建议加上 H1 标签，有些建议则没什么意义，甚至可能有害。掌握基本的优化技术后，自己的想法与软件的建议可以互相印证一下，如果有差异，还是应该以自己的想法为准。

3．内容生成工具

给定关键词，软件自动生成网页内容。可以想象，这种软件生成的内容要么是可读性很低的近乎胡言乱语，要么是自动抓取搜索引擎搜索结果或者其他网站上的内容。不建议使用这种软件，除了用户体验很差外，既可能侵犯他人版权，也可能有复制内容问题。

4．链接生成软件

主要是留言本、论坛、博客评论的群发。这种软件目前在黑帽中很流行，不建议使用。有的搜索引擎对垃圾留言的判断已经相当准确，会把这种链接的权重传递降为零，更严重的是可能会对网站进行某种程度的惩罚。随着所有搜索引擎对垃圾链接判断力的提高，使用链接群发软件将越来越危险。最危险的是，一旦群发，证据就留在其他网站上了，很可能永远无法删除，这就是个随时可能引爆、自己还无法拆除的炸弹。

下面介绍笔者觉得较有价值、经常使用的 SEO 工具。

11.1 Xenu

Xenu 是一款功能简单、但对 SEO 十分重要的蜘蛛爬行模拟工具。它是英文软件。

其界面非常简洁，用户输入一个网址，通常是网站的首页，单击“OK”按钮，这个软件就从所输入的网址顺着链接爬行到其他网址。所以使用 Xenu 可以方便地检查网站内部链接可爬行性，是否有错误链接等，如图 11-1 所示。

Xenu 界面有几个可选的参数。左上角的选择框（Check external links）是让用户选择是否检查外部链接，如果只是想检查本网站链接，则不用选择。

下面有两个文本框，第一个在爬行时把某种特定 URL 当做是内部链接（Consider URLs beginning with this as 'internal'）。比如你有几个网站链接在一起，从输入的第一个 URL 开始，可以同时爬行检测到其他网站。第二个文字框是排除某些 URL（Do not check any URLs beginning with this），比如网站上一些功能链接、带脚本的链接等。

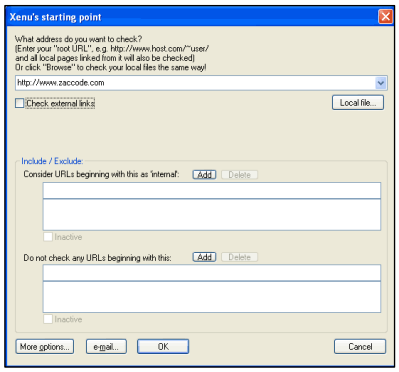


图 11-1　Xenu 软件只需要非常简单的设置

Xenu 运行之后，在结果中就可以看到哪些链接是有错误的，如图 11-2 中标为 not found 的几处。

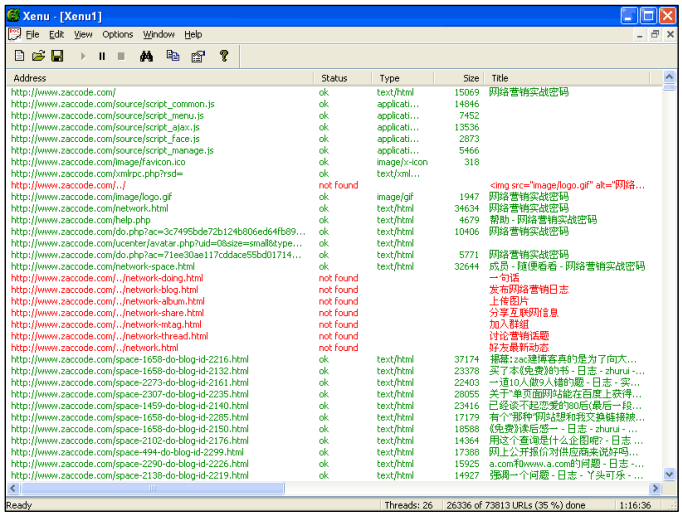


图 11-2　Xenu 软件运行后显示的有问题链接

再认真的站长也可能会发生错误，人工很难检查到这些错链，软件却可以轻易发现。

Xenu 还可以发现一些你并不想让搜索引擎发现和爬行的链接，如图 11-3 中这些动态 URL，是网站上的功能链接，用来发表评论。

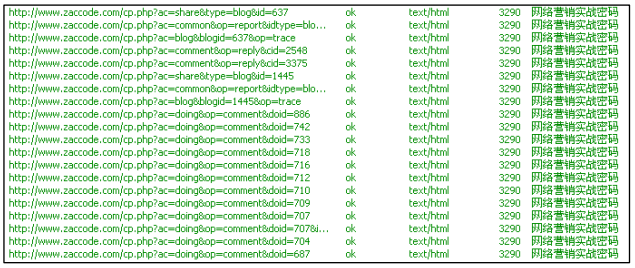


图 11-3　动态 URL

这些链接被收录也没什么实际价值。人工看页面时不容易全都注意到，用 Xenu 运行一遍，就可能看到一些奇怪的、你并不想收录的 URL，然后再通过页面链接调整或直接用 robots 文件禁止搜索引擎抓取。

Xenu 运行完毕后，会把所有错误链接列表供你参考，并且生成一个网站地图。还有一个很有用处的功能是，Xenu 可以按抓取的 URL 生成 XML 网站地图。

Xenu 是一个检查网站内部结构和链接的常用有效工具。下载地址：

http://home.snafu.de/tilman/xenulink.html

11.2 Alexa

Alexa 是最老牌的网站流量显示和排名服务，亚马逊书店旗下网站。其最初数据来源是装有 Alexa 工具条的用户，近几年 Alexa 又引入了其他方面的数据，比如购买网络接入商 ISP 的数据，并且加大了过滤作弊的力度。

网站地址：http://www.alexa.com/

做 SEO 和网络营销的人都知道，所谓“刷流量”是一个产业。很多网站通过作弊手法把 Alexa 数字刷上去，然后去忽悠投资人。这几年 Alexa 经常把大量网站流量数据下调，尤其是中国的网站，通常认为就是在对付“刷流量”作弊。

就算没有作弊因素，Alexa 数据也并不准确，排名越靠后越不准确，原因在于安装 Alexa 工具条的只占全体网民的很小一部分，而且常常是特定的一部分，比如需要查看 Alexa 数据的站长们。

不过无论怎样，Alexa 是一个查看竞争对手网站流量的不错的工具，目前也没有更好的替代品。如图 11-4 所示为点石互动（www.dunsh.org）的 Alexa 数据。

图 11-4 Alexa 排名

最常用的数据是网站流量的世界排名（Traffic Rank）。数字越小，显示流量越大。如同图 11-4 中所显示的，点石网站流量在全世界排名一万名左右，在中国排名 540。这个排名相信是很不准确的，点石网站不可能是全世界前一万名流量大的网站。上点石网站的用户大多是 SEO 人员，很多人装了 Alexa 工具条，导致这个数字有很大失真。

Alexa 也显示流量来自于哪些国家，如点石网站流量 99%来自中国，如图 11-5 所示。

除了流量，Alexa 最近也加了一些有趣的功能，比如显示哪些关键词给网站带来了流量，如图 11-6 所示。

图 11-5　Alexa 显示流量来自哪些国家

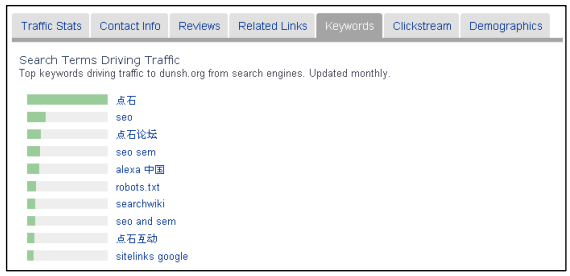


图 11-6　Alexa 显示哪些关键词带来流量

这个数据与真实流量情况接近，但并不准确。

另外一个有意思的数据是点击流（Clickstream），也就是显示出用户在访问点石网站之前（Upstream Sites）和之后（Downstream Sites），各访问了哪些网站，如图 11-7 所示。

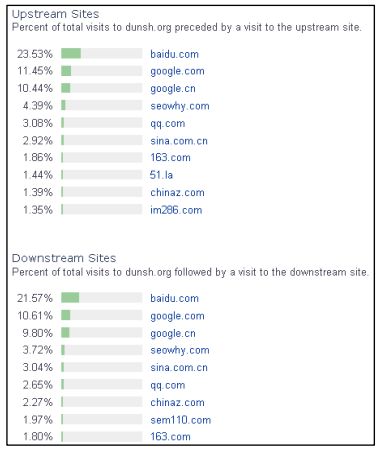


图 11-7　Alexa 显示的点击流

从图 11-7 中可以看到，访问点石的用户还经常同时访问搜索引擎及其他站长相关网站。

Alexa 数据也包括了用户特征统计分布（Demographics），如图 11-8 所示的年龄、学历、性别、上网地点、是否有孩子等数据。

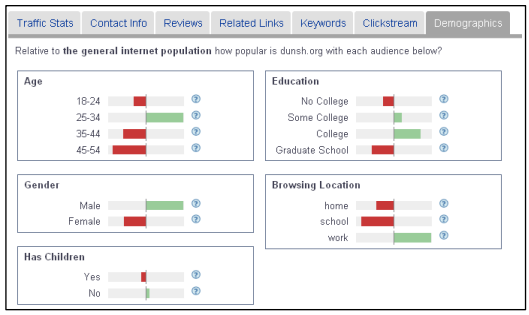


图 11-8　Alexa 显示的用户统计特征

这些数据是否准确无法验证。从几个网站的数据看，有一定的道理，不过只能作为参考。

Alexa 也可以同时显示多个网站的流量比较。如图 11-9 所示是点石互动、本人博客（www.chinamyhosting.com/seoblog/）及网络营销实战密码读者交流区（www.zaccode.com）三个网站的流量对比。

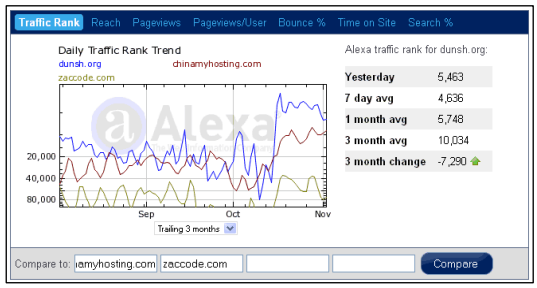


图 11-9　在 Alexa 比较多个网站排名

显然点石流量要大得多，网络营销实战密码流量最小。再看一下平均用户访问页面数（Pageviews/User），如图 11-10 所示。可以看到，点石和网络营销实战密码交流网站的黏度比笔者博客高得多，因为这两个都是互动较多的社区网站。

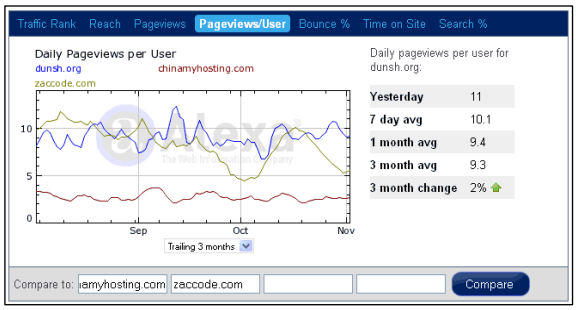


图 11-10　Alexa 显示的用户平均访问页面数

除了流量世界排名，Alexa 还显示网站用户数占全体互联网使用人数之比（Reach）、页面访问量（Pageviews）、平均用户页面访问数（Pageviews/User）、弹出率（Bounce %）、网站访问时间（Time on Site）和来自搜索的流量比例（Search %）。

2009 年 10 月，Alexa 开通了中文版本，网址为 http://cn.alexa.com/，如图 11-11 所示。

图 11-11　中文 Alexa 界面

Alexa 中文版显示数据与英文版完全一样，不过没有显示评论、相关关键词信息等几项内容。可惜，2011 年 12 月初 Alexa 中文版已关闭，访问上面中文版网址会被转向到英文版。

Alexa 是进行竞争对手调查研究时可以参考、但不可全部相信的数据。

11.3　谷歌趋势

谷歌趋势（Google Trends）又分两个功能，一是查看关键词在 Google 的搜索次数及变化趋势，二是查看网站流量（Google Trends for Websites）。

我们先看关键词搜索趋势部分，网址：http://www.google.cn/trends。如图 11-12 所示，Google 趋势用图表形式直观显示关键词搜索量大小及随时间的变化趋势。

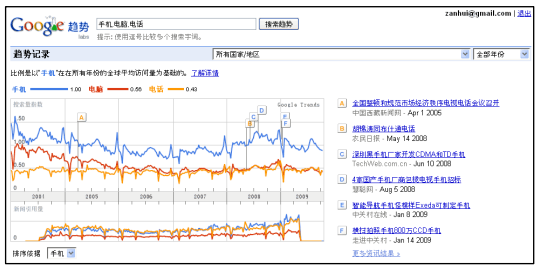


图 11-12　Google Trends 显示关键词搜索量变化趋势

最大的缺点是并没有显示出具体搜索量，而只是给出一个相对数字，用于不同关键词之间的搜索量比较已经足够了。SEO 人员可以使用 Google 趋势进行市场和关键词调研。图 11-12 显示的不同行业关键词搜索量，可能决定了你要进入哪个行业。

Google 趋势可以让用户选择，把数据限定在某个国家和某段时间之内，如图 11-12 中右上角所示。

还可以通过谷歌趋势进行更细致的分析。比如检查同一个产品的不同说法哪一个搜索量更大，以便确定网站主关键词。例如“减肥”和“瘦身”这两个词，如图 11-13 所示。

两个词说的是同一件事，搜索量有巨大不同，搜索“减肥”的远远多于搜索“瘦身”的。这就决定了如果你要做减肥瘦身网站，在不考虑竞争程度的情况下，要想达到最大搜索流量，以减肥为核心关键词更适合。还有很多类似情况，比如关键词是用“电脑”还是用“计算机”？是用“主机”还是用“服务器”？

SEO 人员也可以从 Google 趋势中看到关键词搜索次数随着时间的变化趋势。如图 11-13 所示，搜索“减肥”和“瘦身”的都从 08 年开始有较多增长，可能生活富裕了，需要减肥的人群也更大了。这是一个有上升趋势的市场，前途很光明。搜索趋势对企业选择产品、站长选择网站主题可能具有决定意义。

图 11-13　使用 Google Trends 比较关键词搜索量

再比如，同样是与网站有关的“SEO”和“网站设计”，如图 11-14 所示，可以清楚地看到，搜索“网站设计”的人数呈逐年下降趋势，而搜索“SEO”的则从 2006 年左右开始处于快速上升中，这两者的搜索次数在 2009 年达到了交叉点，很可能今后 SEO 的需求会超过网站设计。这两个行业哪一个在未来会更受关注，前途怎样，一目了然。

Google 趋势还可以显示一些周期性的搜索变化，有利于企业和站长提前准备网站排名，链接交换，甚至人员和产品的准备。如图 11-15 所示。

图 11-14　使用 Google Trends 比较关键词搜索量变化趋势

图 11-15　关键词搜索量周期性波动

有些关键词的周期性升降是显而易见、可以预期的。像搜索“月饼”的，在中秋节时有爆发式增长，中秋节一过，搜索次数快速下降。“粽子”这个词也类似。

有些关键词的周期性升降就不一定那么显而易见了。比如搜索“巧克力”，很明显 2 月 14 日情人节前是搜索巧克力的高峰期。但是在 2008 年 12 月初，为什么又有一个高峰期，甚至这个高峰期的幅度超过了情人节？如果显示全部年份，可以看到每年的 12 月初都有一个高峰，这是为什么？

也许对巧克力行业的人员来说这是完全正常的，但对我来说还真不能预期。一个比较顺理成章的猜测是为了圣诞节及元旦购买礼物，但为什么高峰期是在 12 月份初，比节日提前这么多天呢？我想不出什么合理解释，只能说是不同节日的购物习惯不同吧。对一个要进入这个行业的人来说，也许这个数字就能帮助你提前计划什么时候开始与此相关的内容布置、链接建设、促销配套等。

当然有些词随时间变化不大，像“手表”，除了也在情人节附近有一个很小的增长之外，需求基本上常年稳定。

Google 趋势的第二个重要功能是显示网站流量，Google Trends for Websites。不过这个功能没有中文版，只能在英文版中查看，网址是 http://trends.google.com/websites。

如果查看时没有登录 Google 账号，则流量不显示具体数字。如果登录了 Google 账号，网站的具体流量数字都可以显示出来。

如图 11-16 所示，不同竞争对手的网站流量高下，甚至具体数字一目了然。

当然，真实网站流量只有企业和站长自己知道。第三方数据，包括 Google Trends 数据，都只是估算和参考。就我所能看到真实流量的一些网站来看，很大部分情况下，尤其是网站使用了 Google 流量统计服务 Google Analytics 时，Google 趋势数字比真实流量少 20%～30%。让人高兴的是，越来越多的网站，甚至很多大型网站都在使用 Google Analytics 流量统计。

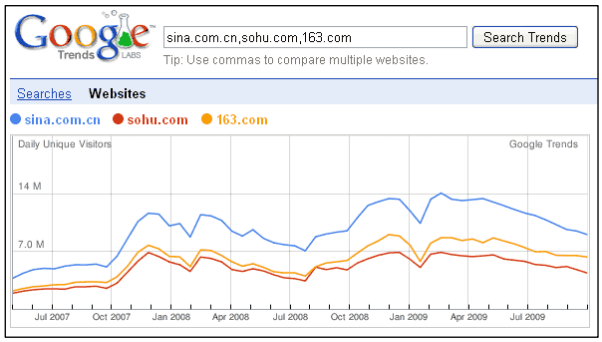


图 11-16　Google Trends 显示网站流量

11.4　百度指数

百度指数是与 Google 趋势非常相似的一个关键词研究工具。如图 11-17 所示，百度指数同样是以图表方式显示关键词搜索量及随时间的变化的。

图 11-17　百度指数显示关键词搜索量变化

比 Google 趋势更好用的是，百度指数标出了与关键词搜索量相关的所谓“用户关注度”，而 Google 趋势只是给出一个与绝对搜索量没有直接关系的相对数值。百度指数中显示的用户关注度数字与搜索次数是什么关系，百度并没有明确说明，SEO 界也众说纷纭。有的人认为用户关注度就是搜索次数，有的人认为不是，而且这两个数值之间相差很大。

我的想法是，用户关注度与搜索次数肯定是相关的，或者说具有比例关系。不同行业和主题可能比例关系不同。对 SEO 人员来说，绝对搜索次数没有有效搜索次数重要。举例来说，搜索“SEO”这个词的绝对数量可能不小，但其中有很大一部分是无效搜索次数，是 SEO 人员在查这个词的排名，却不会点击任何一个结果页面。而搜索“维生素 C”或“红烧肉做法”这种词的，相信绝大部分是普通用户，他们搜索之后，总是要点击结果的，几乎百分之百都是有效搜索。

所以不同行业用户关注度与有效搜索次数之间的比例，很可能是不同的。SEO 人员如果只是为了比较不同关键词之间的相对搜索量，或者观察某个关键词随时间的变化趋势，其实并不用关注具体数字。

如果要预估流量，可以选择一些自己网站已经排到前面的关键词，记录下这些关键词的用户关注度，以及自己网站上这些关键词的百度搜索流量，再加上前面介绍的前 10 名搜索结果点击率，就可以估算出用户关注度与搜索次数的比例关系。这对大型网站来说尤其有效，当掌握的已知排名关键词达到几百几千时，这个比例关系就会趋于准确。

除了关键词搜索趋势及用户关注度数值外，百度指数还列出相关搜索词，以及上升最快相关搜索词，如图 11-18 所示。

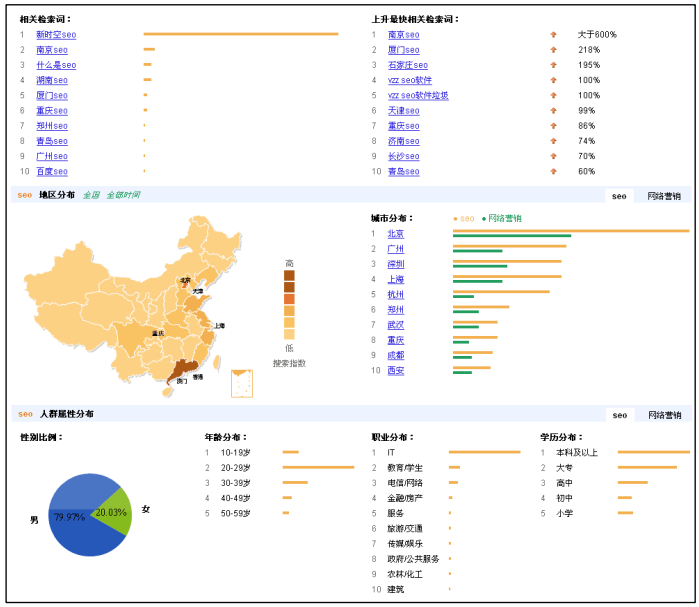


图 11-18　百度指数显示其他相关信息

这对关键词扩展和发现新的有潜力的关键词，也都有很大帮助。

百度指数也列出了用户人群分布数据，如性别、年龄、职业、学历，和 Alexa 数据一样，可作为参考，但不可全部相信。

和 Google 趋势一样，百度指数对周期性关键词的搜索量变化也有清楚的显示，如图 11-19 所示。

使用百度指数的基本版本不需要登录任何账号，网址是 http://index.baidu.com/，百度还给一部分用户提供百度指数账号，使用百度指数账号得到的数据与基本版本是一样的，区别是百度指数账号使用者可以存储关注关键词，建立榜单，如图 11-20 所示。

图 11-19　百度指数显示关键词搜索量周期性波动

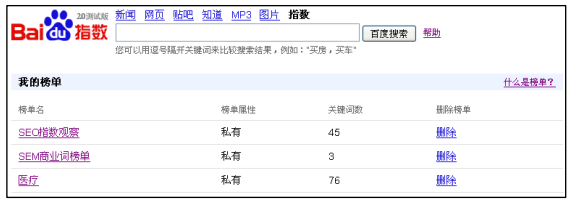


图 11-20　百度指数账号

11.5　百度搜索风云榜

百度搜索风云榜列出主要行业和主题的热门搜索词，以及上升最快的搜索词。这对发现新的有潜力的关键词非常有帮助，如图 11-21 所示。

一些热门搜索词可能几个月甚至几年都保持热门，当然竞争也保持很高水平。

但一些突发的社会热点谁也无法预期。当这些热点发生时，就会反映在上升最快的搜索词中，这就给所有网站一个相同的机会。很多新闻、综合门户类的网站都有 SEO 或栏目编辑人员时刻关注着上升最快的关键词，从中捕捉有潜力的流量来源，一旦发现与自己网站相关的热词，就迅速组织专题，推出大量相关内容。谁先推出内容，谁就将在热点搜索爆发时排名靠前，获得大量流量。这种方法近几年显得非常有效。

图 11-21　百度搜索风云榜

当然，对纯粹电子商务网站来说，这种方法不一定适合。与自己产品无关的流量再多，转化率也会极低，对企业和站长的收入不会有什么影响。

百度搜索风云榜网址是 http://top.baidu.com/，与此类似的工具还有 Google 热榜和搜狗热搜榜，如图 11-22、图 11-23 所示。Google 热榜网址是 http://www.google.cn/rebang/home（可惜目前谷歌热榜已停止服务）。搜狗热搜榜网址是 http://top.sogou.com/。

图 11-22　Google 热榜

图 11-23　搜狗热搜榜

11.6 Google Adwords 关键词工具

Google Adwords 是最重要的 SEO 工具之一，也是笔者最常用的工具之一。

Google Adwords 关键词工具本来是提供给 Google Adwords 广告商扩展、挑选关键词时使用的工具。不过不管是否登录 Google 账号，所有访问者都可以使用这个工具，没有任何限制。如图 11-24 所示，用户只要输入一个关键词，然后单击“获取关键字提示”按钮，这个工具就会列出与此相关的关键词，以及竞争程度和搜索量等数据。

图 11-24　Google 关键词工具

你也可以把数据限制在某种语言和国家。比如只想研究在中国境内搜索简体中文的用户搜索情况，如图 11-25 所示。

图 11-25　在 Google 关键词工具设定语言和国家

单击“获取关键字提示”按钮之后，工具列出大量相关关键字，以及很多重要信息，如图 11-26 所示。

“估算平均每次点击费用”和“广告客户竞争程度”，这两个参数都可以帮助 SEO 人

员估计关键词的竞争程度，这两者之间通常也是成正比的。广告客户竞争程度其实就是参与竞价的广告商数量，一般来说广告商越多，竞价也就越高。但也有特例，比如图 11-26 中所显示的“减肥腹部”这个关键字，虽然参加竞价的广告商数量不是最多，但点击价格却是最高的，而且高得出奇。这显示有可能几个广告商在为争夺第一位血腥拼杀。

图 11-26　Google 关键词工具生成大量相关关键词

“本地搜索量”及“全球每月搜索量”无须解释。值得一提的是，这个搜索量是否是真实数据在 SEO 界也有一些争论。但通常大家认为所显示的搜索量还比较符合事实。根据我自己对一些网站的估算，Google Adwords 关键词工具所显示的搜索量常常超出真实搜索量 30%～40%。

“搜索量趋势”也无须多做解释。这里的趋势比 Google 趋势里所显示的就要粗糙多了，几乎没有太大作用。

Google Adwords 关键词工具里常被忽略却非常重要的一个选项功能是匹配类型。工具提供了 4 种匹配类型：广泛匹配，词组匹配，完全匹配和否定匹配。因为这个工具是给 Google Adwords 广告商使用的，所以匹配类型的用语也源自于 Google Adwords 广告系统。

广泛匹配指的是只要关键词或者相似的词出现在用户的搜索中，就会触发广告。这包括同义词、近义词、单复数、不同的词序等。也就是说如果我们在这个工具中选择广泛匹配，那么列在“减肥”这个词下面的搜索量就包括了减肥、减肥方法、快速减肥、腹部减肥、预测减肥、减肥茶、瘦身等形式。所以这个广泛匹配搜索量其实并不仅仅是“减肥”这个词的搜索量。

词组匹配指的是只要用户搜索的词包含指定关键词，就将触发广告。所以如果在关键词工具中选择词组匹配，列在“减肥”之下的词组匹配搜索量就包括减肥、减肥方法、腹部减肥、快速减肥等词，但不会包括瘦身、塑身这些词的搜索量。词组匹配与广泛匹配类似，但广泛匹配还可能包括瘦身、塑身等同义词。所以词组匹配搜索量也大大超过特定关键词的真实搜索量。

完全匹配才是使用这个工具时通常应该选择的匹配类型。列在“减肥”这个词下的完全匹配搜索量，就真的是搜索“减肥”这个词的次数，这才是 SEO 人员需要准确掌握的搜索量。工具的默认选择是广泛匹配，SEO 人员使用时需要记得选择为完全匹配。

Google 关键词工具是一个非常强大的关键词扩展及预估搜索次数和流量的工具。每搜一个关键词，都会看到一些你原来没想到的相关扩展词。再搜索这些扩展词，又会得到更多的相关搜索，很容易就能够找出成千上万的搜索词，而且能够比较准确地知道搜索量。

如果查询时登录了 Google Adwords 账号，可以直接把你满意的关键词加入到 Adwords 账号中。也可以把这些关键词输出为 Excel 文件，再自行进行大规模整理。

Google Adwords 关键词工具 2010 年做了改版，界面与书中抓图有所不同，但功能一样。新工具界面网址是：https://adwords.google.cn/o/Targeting/Explorer?__c=1000000000&__u=1000000000&ideaRequestType=KEYWORD_IDEAS#search.none。

11.7 微软广告工具

微软广告实验室也提供一套与 SEO 相关的工具。这些工具都不支持中文（不仅没有中文界面，也无法查询中文词或页面），但其中有几个功能十分有特色，在其他地方找不到，所以还是简单介绍一下，有兴趣的读者可以仔细研究。网址是 http://adlab.msn.com/alltools.aspx。这套工具分为几类，包括用户参数数据、关键词研究工具、内容分析工具、视频及互动媒体工具。

比较有特色的一个是商业意图（Commercial Intention）检测工具，可以告诉你某一个关键词的商业性有多高，商业性越高，越容易转化为购买。比如搜索“nokia mobile phone”（诺基亚手机），商业意图程度是 0.93，说明搜索这个词的人有比较高的购物可能性，如图 11-27 所示。

图 11-27　MSN 工具显示的“nokia mobile phone”商业意图程度

而“Blog SEO”的商业意图程度是 0.62，搜索这个词的人购买任何东西的可能性就差得很远了，如图 11-28 所示。当然，这个商业意图度，或者说购买可能性，只能作为参考，就很多关键词来看大致准确。

图 11-28　MSN 工具显示的“blog seo”商业意图程度

另一个很有特色的功能是搜索通道（Search Funnel），这个工具显示用户搜索某一个特定关键词之前和之后，还搜索了哪些关键词。如图 11-29 所示，搜索“China”的用户，在之前还有很多人搜索了：

- Japan，占 1.2%
- India，占 0.77%
- Chinese，占 0.73%
- Hongkong，占 0.68%

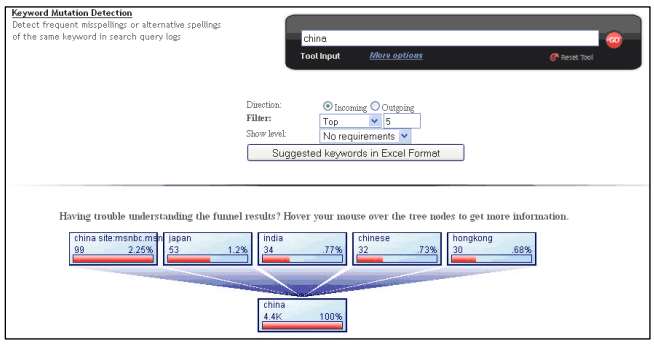


图 11-29　搜索“China”之前还搜索了什么词

而在搜索 China 之后，又搜索了（如图 11-30 所示）：

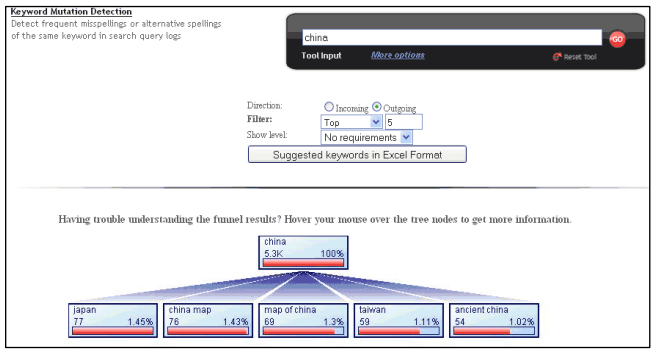


图 11-30　搜索“China”之后又搜索了什么词

- Japan，占 1.45%
- China map，占 1.43%
- map of China，占 1.3%
- Taiwan，占 1.11%
- ancient China，占 1.02%

这对我们了解用户搜索意图很有帮助，因为搜索通道显示了用户不断细化和修改搜索的过程，体现出他到底想要什么。如图 11-30 中所显示的，一部分人搜索“China”之后，没有找到需要的结果，转而继续搜索“中国地图”，一部分人继续搜索“古代中国”。这告诉我们搜索“中国”这个宽泛词的用户，其实上是想寻找什么具体信息。

前面提到过搜索词有不同种类：信息型、导航型和交易型。这个工具可以帮助我们了解，用户在搜索过程中，是怎样从信息型转变为导航型和交易型的。网络营销人员可以据此修改网页，提高转化率，捕捉更好的目标关键词。

11.8 Google 百宝箱

Google 百宝箱（Search Options）是 Google 在 2009 年推出的功能，其中包含几个对 SEO 有益的信息。用户搜索某个关键词后，在页面上部会看到“打开百宝箱”链接，如图 11-31 所示。

图 11-31 Google 百宝箱

单击“打开百宝箱”链接后页面左侧就会出现百宝箱选项，如显示所有结果或视频、论坛结果，按时间显示结果，排序方式变化，视图变化，以及显示格式的变化，如图 11-32 所示。

其中按时间显示功能可以帮助 SEO 人员检查网站收录情况。比如搜索“site：域名”，然后选择“一周内”结果，所显示的就是这个网站一周内收录的新页面，如图 11-33 所示。

图 11-32　Google 百宝箱选项

图 11-33　一周内新收录的页面

前面提到过，对大中型网站来说，解决收录问题是排名和流量的基础。百宝箱可以让你轻松查看新页面收录数量和速度。除了按一小时内、一天内、一周内，甚至一年内显示，还可以自定义日期范围，非常方便。百度高级搜索也与此类似，如图 11-34 所示。

图 11-34　百度高级搜索也可以限定时间

除了选择搜索结果抓取时间，百度高级搜索还可以选择文件类型等。网址是 http://www.baidu.com/gaoji/advanced.html。

单击 Google 百宝箱“图文并茂”格式显示，将会看到页面说明右侧加上了 Google 从页面上抓到的图片，这是个有意思的功能，如图 11-35 所示。

图 11-35　Google 百宝箱“图文并茂”格式

另外一个对 SEO 有帮助的功能是神奇罗盘（Wonder Wheel）。搜索一个特定关键词后，神奇罗盘将会显示出相关关键词，其内容与通常显示在页面底部的相关搜索相同，如图 11-36 所示。但是大家可以看到，以神奇罗盘格式显示相关关键词非常快速、直观、有趣。点击任何一个相关关键词后，又会自动生成更多相关的关键词。可以一直单击下去，观察相关的关键词链条。这对 SEO 人员快速扩展关键词十分轻松有效。

Google 不停地对搜索结果页面进行优化和改动。2010 年后，百宝箱中的内容成为默认设置，用户搜索关键词后，结果页面左侧自动显示原来百宝箱中的选项，页面左上角的“打开百宝箱”当然也取消了。如图 11-37 所示。

2011 年 7 月，Google 取消了很多人喜欢的神奇罗盘功能。

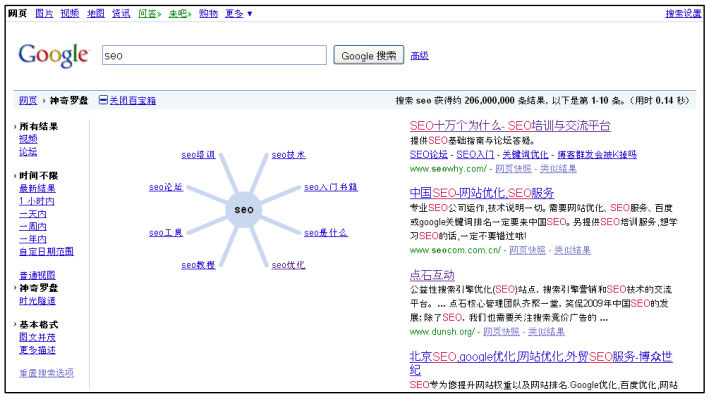


图 11-36　Google 神奇罗盘

图 11-37 原来“百宝箱”中的内容成为 Google 结果页面左侧常规选项

11.9 Google 快讯

Google 快讯（Google alerts）也是我个人非常喜欢的一个工具。Google 快讯不仅对 SEO 有用，对网上公关危机处理更有用。

网址是：http://www.google.com/alerts?gl=cn&hl=zh-CN。

用户只要填写想监控和关注的关键词，Google 会自动通知你与这个关键词相关的最新内容。如图 11-38 所示。

图 11-38 Google 快讯

除了搜索词之外，用户可以设定内容类型，包括“新闻”、“在线文章”、“论坛”或“全部”，通常要监控全部。还可以设定通知频率，包括一天一次，一周一次和出现新的结果时。通常我喜欢设定为出现新的结果时。

另一个选项是发送到什么地方，你可以选择发送到你的邮箱，或者生成一个 Feed，然后可以像订阅任何普通 Feed 如博客一样，在你喜欢的 Feed 阅读器中每天查看。比如

你想监控网上有什么关于你公司的评论，就可以把搜索词设定为公司名称，频率设定为出现新的结果时。这样每当网上（当然，仅限于 Google 收录的那部分）有内容提到你的公司名称，你都将收到通知，赶紧去看一下人家在讨论你什么。

你也可以监控网站某个新栏目页面收录进度，只要把搜索词设定为“site:domain/category/”，其中 domain 是你的域名，category 是栏目目录名。这样，每当有新页面被收录时，你都会接到通知。你也可以使用 Google 快讯监控你自己网站或者竞争对手网站的新链接，只要把搜索词设定为“link:domain”。

11.10 服务器头信息检测器

前面多次提到服务器状态码，如 301 转向、302 转向、404 错误页面。

很多域名提供商允许用户做域名转发。但这个“域名转发”可以用很多不同的机制实现，是一个很含糊的概念，既可以是 301 实现转发，也可能是 302，还可能是 JS 脚本，也可能是 Meta 刷新，甚至可能是 Flash 实现的。

使用哪种机制实现转向，用户在浏览器上访问页面并不能看出来，因为结果都是立即转到另外一个 URL。要判断是哪种机制实现转向，就需要检查服务器头信息。

网上有很多检测服务器头信息的工具。我个人比较喜欢用的是：

http://www.webrankinfo.com/english/tools/server-header.php，如图 11-39 所示。

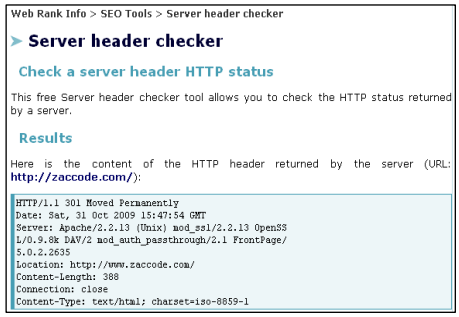


图 11-39 服务器头信息检测显示的 301 转向

检查 URL http://zaccode.com/ 的服务器头信息，从显示的结果可以看到，这是一个 301 转向，指向 http://www.zaccode.com/，说明设置正确，目的当然就是预防或解决网址规范化问题。同时显示的还有服务器的一些信息。

如果检查网址 http://www.zaccode.com/，如图 11-40 所示，得到的就是一个正常的 200 状态码。

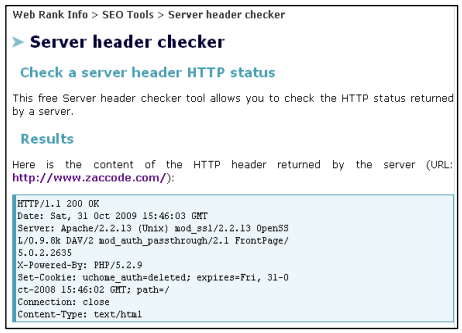


图 11-40　服务器头信息检测显示的 200 状态码

有一些服务器设置有问题，在需要显示 404 状态码时，却返回 200 状态码。虽然用户看到的都是页面没有找到之类的错误信息，但不同状态码有天壤之别，将使搜索引擎做不同的处理。返回 200 状态，将使搜索引擎认为你的网站出现很多重复内容。而返回 404 状态码，将使搜索引擎知道这些页面不再存在，会逐渐删除。

11.11　W3C 验证

通过 W3C 验证与否与搜索引擎排名是否有关系，SEO 行业有不同看法。大部分人倾向于认为是否通过 W3C 验证与排名没有关系，或至少影响极小。Google 反作弊组工程师 Matt Cutts 也曾经说过，W3C 验证并不是排名算法的一部分，因为搜索引擎知道大部分网页是通不过 W3C 验证的，都包含有代码错误。

就我个人的观察，很少看到完全通过 W3C 验证的页面。这很少一部分能通过验证的页面，排名并没有看到有什么不同。找几组关键词查看，排在前面的页面检查 W3C 验证后就会发现，两者之间没有关系。错误比较多的页面，并不比错误少的页面排名差。

当然，W3C 毕竟是一个标准，如果能通过自然最好。验证 W3C 标准，最权威的莫过于 W3C 官方验证工具：http://validator.w3.org/ 。

从图 11-41 可以看到，百度和 Google 首页这两个相当简洁的页面，也都有 40 个左右错误及警告。

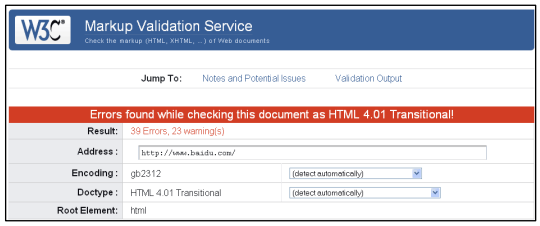


图 11-41　W3C 验证百度首页

如果不能通过 W3C 验证会影响搜索引擎排名，相信搜索引擎本身会把页面代码先通过验证。所以能通过验证当然最好，不能通过也大可不必担心。

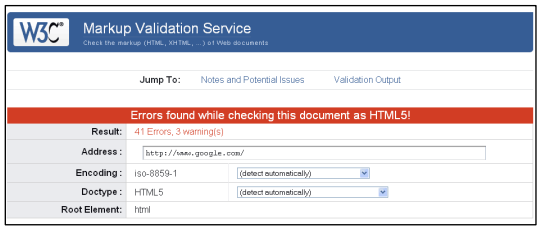


图 11-42　W3C 验证 Google 首页

代码不能通过验证可能会影响搜索引擎排名的唯一情况是，代码中的严重错误导致搜索引擎无法解析页面，无法读取文字内容，或者无法跟踪页面中的链接。这就将造成网站收录和相关性判断方面的问题。

11.12　雅虎外链检查工具

如果要给所有 SEO 工具做个排名，雅虎外链检查工具应该排名第一，原因是我们找不到其他工具能显示搜索引擎数据库中比较真实的链接数据。

非常可惜的是，雅虎和微软于 2009 年 8 月达成协议，雅虎停止自己的搜索技术，转而使用微软的搜索技术和数据库。雅虎搜索使用微软技术已于 2010 年开始实施，2011 年初完成。这个深受 SEO 喜爱、唯一能查看搜索引擎真实数据的工具也于 2011 年 11 月 21 日被关闭。虽然这个工具已不存在，但本节保留，没有机会用到这个曾经最重要的 SEO 工具的新人们可以感受一下它的重要意义。

雅虎外链检查工具正式名称是 Yahoo Site Explorer，网址是：http://siteexplorer.search.yahoo.com/。

链接建设是 SEO 最重要的工作之一。检查自己网站和竞争对手网站的外部链接，是判断网站优化难度、发现新链接来源的重要手段。而外部链接数据只有搜索引擎才能提供。Google 和微软必应的 link:指令也可以用来查询链接，不过所显示的链接只是搜索引擎数据库中的一部分，而且显示哪些、按什么顺序显示，没有任何规律，基本上不能作为判断外部链接的依据。只有雅虎外链工具才显示比较全面的链接数据，而且大致上是按照链接的重要性排序的。

在雅虎 Site Explorer 输入一个 URL 后，单击“Explore URL”按钮，雅虎将返回两种数据，如图 11-43 所示。

单击结果页面上部的“Pages”按钮，显示的是雅虎收录的网站页面数。这个数据虽然也有一定意义，但大部分人使用 Yahoo Site Explorer 不是为了查看这个。

第二个按钮“Inlinks”，单击后显示的就是链接数据，这是最重要的，在其他地方得不到的数据。链接数据又有两个可选项。第一个下拉菜单有三个选择：

- 显示来自所有页面的链接（From All Pages）。
- 不显示来自本域名的链接（Except from this domain）。
- 不显示来自本子域名的链接（Except from this subdomain）。

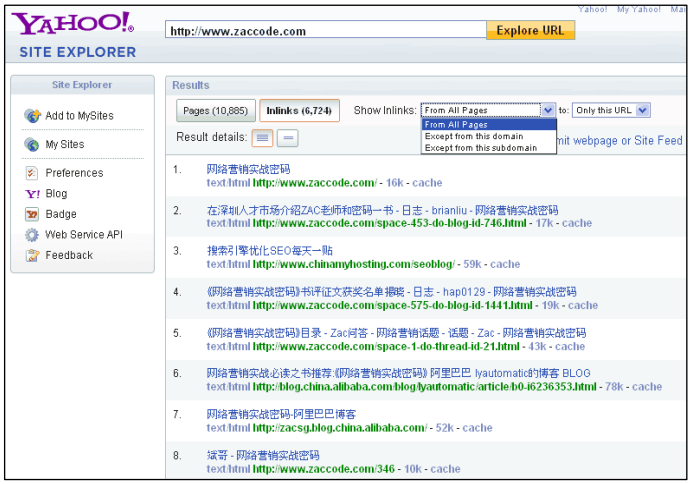


图 11-43　Yahoo! Site Explorer 显示的页面全部反向链接

所以如果想查看网站的外部链接，需要选（2）或（3）。

第二个下拉菜单有两个选项：

（1）只显示连到这个 URL 的链接（Only this URL）。

（2）显示连到整个网站的链接（Entire Site）。

所以这个工具既可以检查整个网站得到的外链，也可以检查某一个特定页面得到的外部链接。如图 11-44 所示，www.zaccode.com 这个网站，首页有 162 个外部链接。

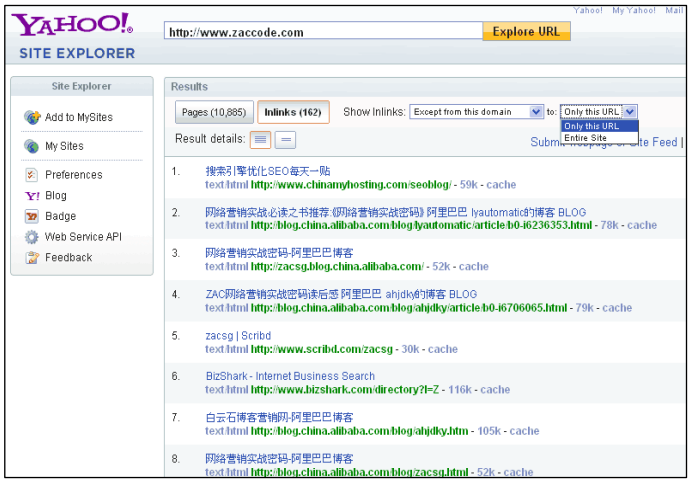


图 11-44　Yahoo! Site Explorer 显示的页面外部链接

雅虎链接检查工具的一个缺点是，会显示带有 nofollow 属性的链接，在结果列表中不加以区分。

另一个缺点是，虽然可以显示外部链接总数，但只显示前 1000 个链接 URL。对外部链接稍微多一点的网站，比如几千个以上，我们只能知道链接总数，却不容易查看到全部链接来自哪些网站和 URL。

中国雅虎也提供了类似工具，名称为“站长工具”，网址是 http://sitemap.cn.yahoo.com/，如图 11-45 所示。

可以看到，中国雅虎站长工具显示的数字与英文雅虎差别比较大。由于中国雅虎的搜索数据库与百度、Google、英文雅虎、微软必应有一定差距，参考价值比英文 Yahoo Site Explorer 小些。

图 11-45　雅虎站长工具

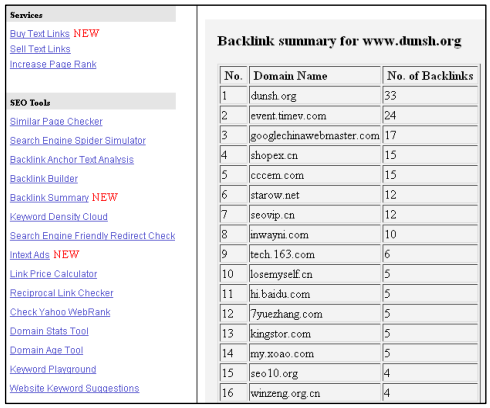
11.13　外链概况工具

用 Yahoo Site Explore 得到的外部链接虽然比较全面，但格式显得比较繁杂。要了解一个网站的外链概况需要翻很多页，还不一定能看清到底有哪些域名链接过来。

如图 11-46 所示，这个线上工具可以显示一个网站的外部链接概况：

http://www.webconfs.com/backlink-summary.php

输入域名后，工具返回一个非常简洁的列表，告诉你有哪些域名提供了外部链接，每个域名给了多少外部链接。虽然数字不一定非常准确，但 SEO 人员可以一目了然地看清外部链接概况。



Services

Buy Text Links NEW
Sell Text Links
Increase Page Rank

SEO Tools

Similar Page Checker
Search Engine Spider Simulator
Backlink Anchor Text Analysis
Backlink Builder
Backlink Summary NEW
Keyword Density Cloud
Search Engine Friendly Redirect Check
Intext Ads NEW
Link Price Calculator
Reciprocal Link Checker
Check Yahoo WebRank
Domain Stats Tool
Domain Age Tool
Keyword Playground
Website Keyword Suggestions

Backlink summary for www.dunsh.org

No.	Domain Name	No. of Backlinks
1	dunsh.org	33
2	event.timev.com	24
3	googlechinawebmaster.com	17
4	shopex.cn	15
5	cccem.com	15
6	starow.net	12
7	seovip.cn	12
8	inwayni.com	10
9	tech.163.com	6
10	losemyself.cn	5
11	hi.baidu.com	5
12	7yuezhang.com	5
13	kingstor.com	5
14	my.xoao.com	5
15	seo10.org	4
16	winzeng.org.cn	4

图 11-46　外链概况

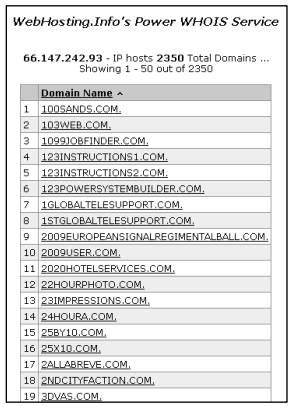
11.14　IP 地址检查工具

很多站长经常担心自己网站所在的服务器或 IP 地址上有其他网站被搜索引擎删除或惩罚，进而影响自己的网站。这种情况并不常见。世界上大部分网站是放在虚拟主机上的，每个 IP 地址对应几十、几百个网站是很正常的事情，这些网站之间并没有联系。搜索引擎也明白这一点，所以不会轻易因为服务器上的一个网站作弊而连带惩罚其他网站。

当然，如果同一台服务器上的大部分网站域名属于同一个人，这些网站又全都作弊，那么这个 IP 地址上的干净网站也可能因此被惩罚。所以检查一个 IP 地址上还放了其他哪些网站，是挑选虚拟主机和检查网站为什么被惩罚时经常需要的数据。

如图 11-47 所示，以下这个工具就可以告诉你一个 IP 地址上都放了哪些网站：

http://whois.webhosting.info/66.147.242.93



WebHosting.Info's Power WHOIS Service

66.147.242.93 - IP hosts **2350** Total Domains ...
Showing 1 - 50 out of 2350

	Domain Name
1	100SANDS.COM.
2	103WEB.COM.
3	1099JOBFINDER.COM.
4	123INSTRUCTIONS1.COM.
5	123INSTRUCTIONS2.COM.
6	123POWERSYSTEMBUILDER.COM.
7	1GLOBALTELESUPPORT.COM.
8	1STGLOBALTELESUPPORT.COM.
9	2009EUROPEANSIGNALREGIMENTALBALL.COM.
10	2009USER.COM.
11	2020HOTELSERVICES.COM.
12	22HOURPHOTO.COM.
13	23IMPRESSIONS.COM.
14	24HOURA.COM.
15	25BY10.COM.
16	25X10.COM.
17	2ALLABREVE.COM.
18	2NDCITYFACTION.COM.
19	3DVAS.COM.

图 11-47　共用同一 IP 地址的网站

只要输入一个 IP 地址，就得到这个 IP 上所有网站的列表。站长可以再随机检查其中一部分网站，看是否有被搜索引擎删除或严重惩罚的现象。如果很多网站都被惩罚，那么这个 IP 地址可能是有问题的，最好避免使用。

11.15 关键词问答

这个工具并不支持中文，不过因为非常独特，而且对英文 SEO 非常有用，所以还是介绍一下。

WordTracker 是一款著名的关键词扩展工具。它有一个免费的线上小工具：

http://labs.wordtracker.com/keyword-questions/

用户输入一个关键词，工具就返回基于这个关键词的一些问答式搜索，如图 11-48 所示。

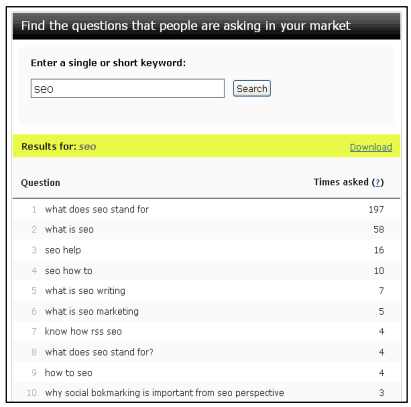


图 11-48　包含关键词的问句式搜索

比如输入“seo”，这个工具就返回下面这些问题式的搜索：

- SEO 代表什么。
- SEO 是什么。
- SEO 帮助。
- 怎样做 SEO。
- 什么是 SEO 文案写作。
- 什么是 SEO 营销。
- 为什么社会化书签从 SEO 角度说很重要。

这类问题都是真实的用户在搜索引擎中搜索的，虽然搜索数量不大，也可以在一定程度上扩展关键词。

对 SEO 来说，这个工具更重要的不是扩展了关键词，而是直接告诉你应该写什么文章。只要拿出这些问句作为文章标题，写上几百字的文章，这个问句式的搜索词就有很大希望得到排名了。

输入与你网站主题相关的不同关键词，可以轻易扩展出几百几千篇很有针对性的文

章。对很多不知道该写什么原创内容的站长来说，这是一个很有帮助的工具。

11.16 Google 搜索解析

Google 搜索解析（Google Insights for Search）与 Google 趋势有点类似，但提供的数据比 Google 趋势更深入。

如图 11-49 所示，用户输入一个或几个关键词，Google 搜索解析将按图表格式给出相对关注度和随时间变化的趋势。

图 11-49　Google 搜索解析

同样，关注度并不是绝对搜索次数，只是标准化了的 1～100 的数值。

按 Google 的说法，工具所显示的是关注度，而不是搜索量。所以甚至一条下降曲线，也并不一定代表搜索量是在减少的，因为关注度是这个关键词相对所有关键词搜索次数的比例。如果整体搜索次数在上涨，那么关注度下降并不一定意味着这个关键词的搜索次数在下降。

同时，搜索解析也显示出这些关键词在不同地域的受关注程度，并列出相关的热门搜索和上升最快的搜索，如图 11-50 所示。

这个工具与 Google 趋势最大的不同在于，界面右上角提供了几个过滤器。

第一个过滤器可以把搜索限定在 Google 的特定垂直服务中，可选择 Google 网页搜索、Google 图片搜索、Google 资讯搜索和 Google 购物搜索。

第二个过滤器是限制地理位置，你可以把搜索趋势限定在某个国家，甚至某个省，如图 11-51 所示。

第三个过滤器是限定时间，你可以研究某一年或某个月数据，也可以自己设定任何

日期范围，如图 11-52 所示。

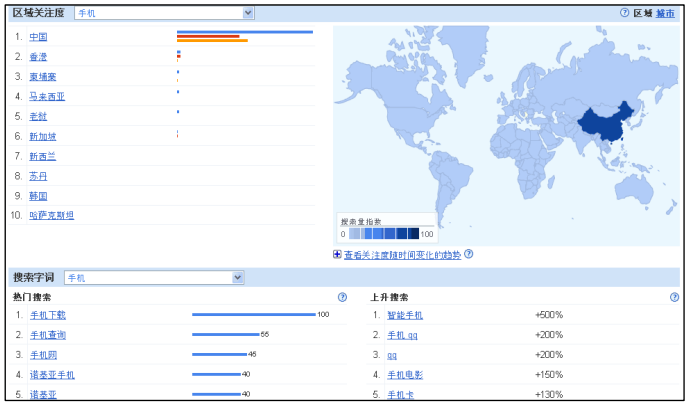


图 11-50　Google 搜索解析的地域关注度

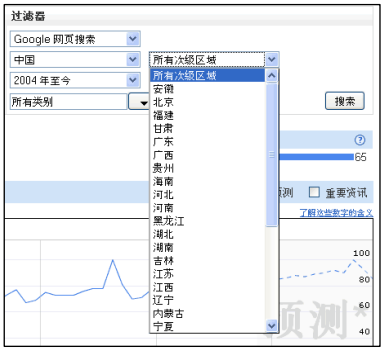


图 11-51　Google 搜索解析的地域过滤

图 11-52　Google 搜索解析的时间过滤

第四个是类别限定，还不支持在中国的数据。

所以，综合使用 Google 搜索解析的不同过滤器，你甚至可以研究广东省 2008 年春节期间搜索智能手机这类非常具体的关键词数据。

Google 搜索解析的网址是 http://www.google.cn/insights/search/。

11.17　网站索引指数查询

网站索引指数查询（IndexRank）这个工具显示网站新页面收录速度和数目。网址是：

http://www.soloseo.com/tools/indexRank.html。

这个工具实际上就是使用 Google 百宝箱限定某段时间的搜索结果，从而得出网站页面收录速度的指数。如在 Google 百宝箱工具介绍中提到的，使用 site: 指令，并把时间限定在两星期、一个月、三个月等，就可以得到 Google 在这段时间内所收录的新页面。网站索引指数查询工具把这些数字整合起来，计算出一个索引指数，0～10，数字越大说明网站被

收录得越快，当然也就意味着网站内部结构比较好，同时有更新内容，如图 11-53 所示。

图 11-53　网站索引指数

不过这个工具也只是作为参考，除了网站本身规模差别很大之外，Google 的 site:指令返回的数字很多时候就不准确。

11.18　SEO for Firefox

SEO for Firefox 是一个综合显示页面 SEO 数据的火狐浏览器插件，是我个人最常用的 SEO 工具之一。

虽然网上有很多网站数据查询工具，包括搜索引擎本身所提供的，但是要针对多个竞争对手检查所有这些数据，也是件很麻烦的事情。SEO for Firefox 这个 Firefox 浏览器插件就集成了大部分 SEO 关心的网站数据，使竞争对手调查变得非常简单。

这个插件是著名的 SEO Book.com 博主 Aaron Wall 开发的，如图 11-54 所示。下载地址是：

http://tools.seobook.com/firefox/seo-for-firefox.html。

图 11-54　SEO for Firefox 火狐插件

安装并启动这个插件后，在 Google 或雅虎搜索任何关键词时，页面除了显示正常的搜索结果外，还在每个结果下面列出了大量 SEO 有关数据，包括：

- 网页 PR 值。
- Google 快照更新日期（Google Cache Date）。快照越新，说明被抓取得越频繁，权重越高。
- 流量价值（Traffic Value）。
- 网站年龄（Age）。这个数据不是从域名注册日期开始算起的，而是从互联网档案馆（后面有介绍）第一次出现网站内容开始算起的。
- 美味书签（del.icio.us）中被收藏次数。 del.icio.us 是书签分享服务，最重要的英文社会化媒体网站之一。
- Digg 中出现的次数。digg 是内容提交、分享网站，也是最重要的英文社会化媒体网站之一。
- Stumbleupon 中出现的次数。Stumbleupon 也是内容提交、分享网站。
- Twitter 中出现的次数。微博网站鼻祖。
- Yahoo 收录的全站链接（Y! Links）。这个工具显示的 Yahoo 链接不是前面介绍的 Yahoo Site Explorer 查询数据，而是使用 linkdomain:指令得到的数据。
- Yahoo 收录的全站.edu 链接（Y! .edu Links）。
- Yahoo 收录的全站.gov 链接（Y! .gov Links）。
- Yahoo 收录的本页面链接（Y! Page Links）。
- Yahoo 收录的本页面.edu 链接（Y! .edu Page Links）。
- Technorati 数据。最重要的博客搜索服务之一。
- Alexa 排名。
- Compete 排名。
- Compete 独立 IP 数。
- Google 趋势链接（Trends）。链接到显示网站流量的 Google 趋势页面。
- Google 快照数量（Cached）。也就是被 Google 收录的页面数。
- 开放目录收录次数（dmoz）。
- Blog lines 中被博客订阅的人数。Blog lines 是常用的 RSS 阅读器之一。
- Google 博客搜索中 link:指令得到的链接数（Page blog links）。
- 维基百科中出现次数（wikipedia）。
- Yahoo 目录收录情况（dir.yahoo.com）。
- BOTW 收录情况（Botw）。BOTW 是权重比较高的网站目录。
- Business 收录情况。Business 也是权重比较高的网站目录。
- 域名注册信息（whois）链接等。

在做关键词和竞争对手研究时，只要搜索一个关键词，排在前面的所有网站数据一目了然。除了网站本身页面优化情况看不到外，网站的链接、权重、社会化营销等数据，都可以在一个页面内轻松对照，排在前面的网站为什么排名好，大致能看出个端倪。

11.19 SEO 工具条

SEO 工具条（SEO Toolbar）也是 Aaron Wall 开发的常用 SEO 火狐插件。下载地址：http://tools.seobook.com/seo-toolbar/。

在火狐浏览器中安装插件后，以工具条形式显示正在访问网页的 SEO 数据，包括 Google PR 值、从 Yahoo 查询的链接数据、开放目录、雅虎目录、BOTW 等目录的收录情况、互联网档案馆所显示的网站年龄、Compete 排名等，如图 11-55 所示。

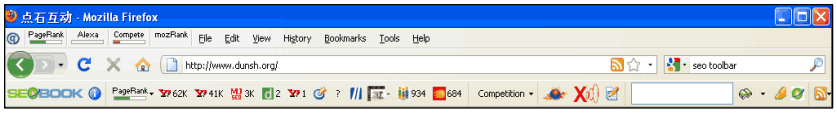


图 11-55　SEO Toolbar 火狐插件

其中一些按钮还可以点击查看更详细的信息。比如最左侧的蓝色带有字母 i 的按钮，点击之后显示当前页面的一些信息，包括 PR 值、搜索引擎收录情况、页面更新日期、雅虎查询得到的各种全站链接和本页面链接、主要目录收录情况、流量数据、社会化媒体数据，如图 11-56 所示。

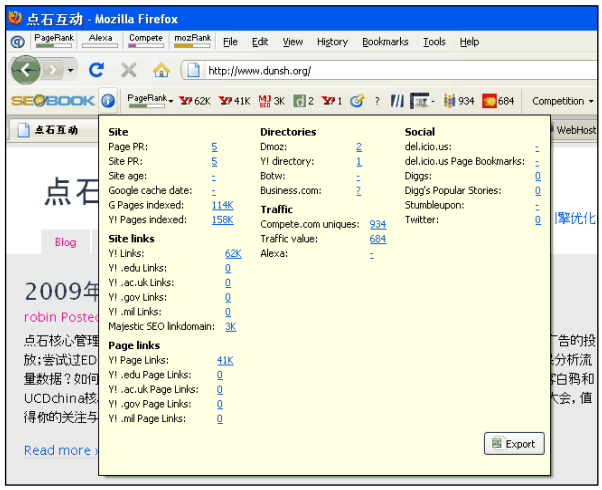


图 11-56　SEO Toolbar 显示当前页面信息

这部分数据与 SEO for Firefox 显示的类似。

工具条中间写有 Competition 的下拉菜单，提供更多 SEO 数据，如图 11-57 所示，包括：

- 本页面 IP 地址。
- 服务器头信息（Server Headers）。
- 更改浏览器类型（User Agent），用来检查网站是否使用了隐藏页面（cloaking）。

- 连往 Compete、Google 趋势、Alexa、Google 基于搜索的关键词等工具的链接。

红色 X 按钮被称为 SEO X 光检查，点击之后显示页面本身的一些 SEO 相关数据，如外部链接、内部链接、总链接数、关键词密度（这个对中文页面不准）、页面标题标签，出现 H 标签的地方也被显示出来，如图 11-58 所示。

图 11-57　SEO Toolbar Competition 菜单里的数据

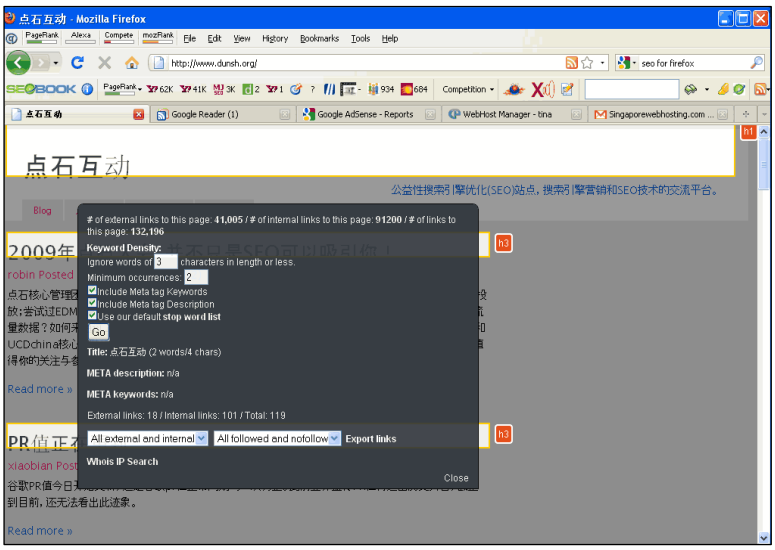


图 11-58　SEO Toolbar SEO X 光检查

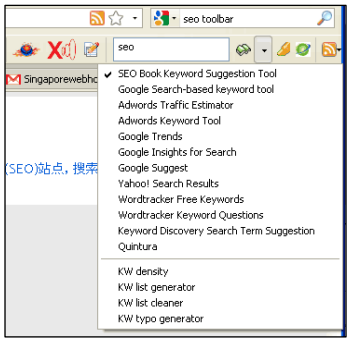
工具条上还有一部分是与本页面无关的通用的关键词研究工具。输入一个关键词后，单击下拉菜单，就可以直接访问其他关键词研究工具，包括 Google 关键词工具、Google 趋势、Google 搜索解析、WordTracker 关键词问答工具等，如图 11-59 所示。

这些工具大部分在本书中都有单独介绍。

11.20　火狐浏览器 Search Status 插件

Search Status 是火狐浏览器使用者最常用的插件之一，用来显示正在访问的页面 PR

值，Alexa 排名，Compete 排名，以及最近刚刚加上的 mozRank，如图 11-60 所示。其中 mozRank 是我使用这个工具的最大原因。



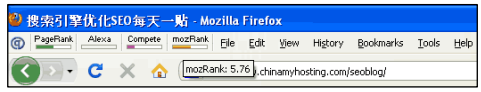


图 11-59　SEO Toolbar 的常用 SEO 工具链接　　图 11-60　Search Status 火狐插件

mozRank 里的 moz 指的是 seomoz.org，世界上最有名的 SEO 博客和服务商之一，相信国内做 SEO 的也都知道。SEOmoz 于 2007 年引入外部投资，开发了一系列 SEO 工具。这套工具最与众不同的地方是基于 SEOmoz 自己索引网上页面而建立的数据库基础之上，类似于搜索引擎自身的页面索引库。

除了工具条 PR 值，搜索引擎很少透露他们自己数据库中的链接数据。SEOmoz 自行建立页面索引库，就可以掌握互联网上大部分主要页面之间的链接，找到链接与排名之间的对应关系。

当然，SEOmoz 的索引库爬行广度、深度都不能与真正的搜索引擎相比，规模要小得多，但重要页面基本上都有收录，所以还是能获得很多其他地方无法看到的数据。

mozRank 就是 SEOmoz 仿照 PR 计算公式，使用自己页面索引库中的链接数据得到的类似于 Google PR 的一个值。mozRank 值比工具条 PR 值显示得更精确，精确到小数点后两位。

一般来说，页面的 Google 工具条 PR 值与 mozRank 值应该相差不多，因为是按同一个公式计算的，只是参与计算的页面数不一样。mozRank 漏掉的大多是不太重要的页面，所以两个结果大致相当。

如果某个页面 PR 值大大低于 mozRank 值，很可能说明这个页面受到了 Google 的某种惩罚，或者某些反向链接因为某种原因被过滤掉，没有计入 PR 值，这都不是好信号。

右键单击 Search Status 插件左侧的@字符按钮，还可以看到一些 SEO 相关工具链接和选项，如图 11-61 所示。

比如：

- 高亮 Nofollow 链接（Highlight Nofollow Links），比较有用，不用再看页面源代码检查 Nofollow 链接。

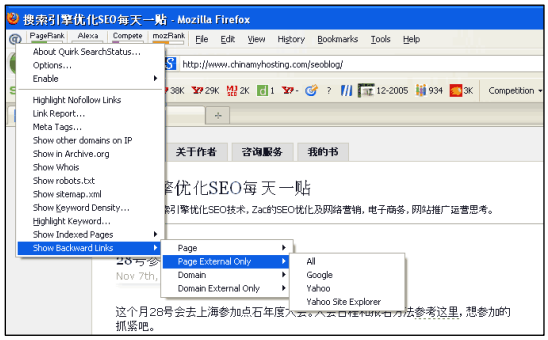


图 11-61　Search Status 火狐插件中的工具链接

- 链接报告。
- meta 标签。
- 同一个 IP 地址上的其他域名。这个链接是指向微软必应的 IP:搜索指令。必应的这个指令会显示在这个 IP 地址上的所有收录页面，包括不同的域名，所以也可以检测到同一个 IP 上有哪些域名。
- 互联网档案馆链接。
- 域名注册信息链接。
- 各主要搜索引擎收录页面。
- 各主要搜索引擎反向链接数据。

下载地址：https://addons.mozilla.org/en-US/firefox/addon/321。

11.21　火狐 SeoQuake 插件

SeoQuake 插件也是一个显示页面 SEO 信息，帮助进行竞争对手调查的强有力工具，和 SEOBook 的 SEO for Firefox 插件类似。

下载地址：http://www.seoquake.com/。

对国内 SEO 来说，最重要的是这个插件支持百度。它的缺点是显示排版格式比较混乱，而且由于查询量比较大，容易被搜索引擎屏蔽使用者的 IP 地址。

SeoQuake 有三个主要显示方式：工具条，页面数据条，以及在搜索页面上显示数据。

如图 11-62 所示，火狐浏览器地址栏下面就是以工具条显示的数据。

与 SEO for Firefox 类似，显示的包括各搜索引擎的收录数、链接数、主要目录收录情况、社会化媒体网站收录情况、域名所在 IP 地址等。

第二个显示方式就是图中页面上部看似工具条，但实际上是以 CSS 控制的加在页面本身上的数据条。它显示的数据与上面的工具条是完全一样的，所以这两个功能视用户的喜好，启动一个就可以了。

图 11-62　SeoQuake 火狐插件

第三种显示方法是在搜索结果页面中显示数据，这个功能与 SEO for Firefox 插件类似，优势是支持百度。如图 11-63 所示是在 Google 搜索结果页面上显示的数据。

图 11-63　SeoQuake 插件显示 Google 搜索结果页面数据

可以看到当显示的数据比较多时，一行显示不下，这个插件使用了框架，要想看全数据，就要左右拉动页面。

如图 11-64 所示是在百度搜索结果页面所显示的数据。

虽然显示排版比较混乱，但因为支持百度，对国内站长来说也很有用。

这个插件的另外一个灵活之处是，可以任意设置显示哪些数据。如图 11-65 所示是插件的设置选项，在工具条、数据条及搜索页面上，你都可以选择要显示哪些数据。

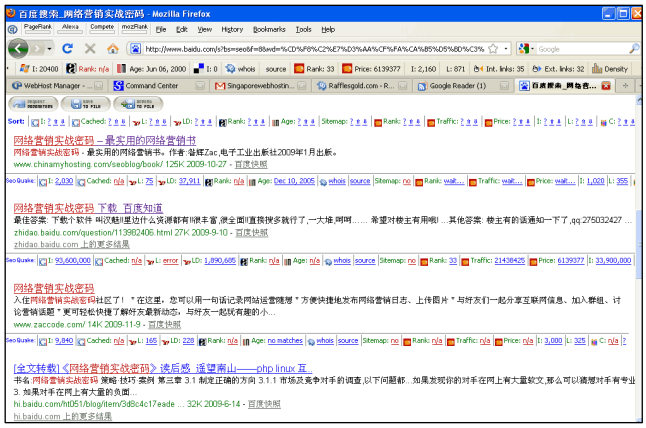


图 11-64　SeoQuake 插件显示百度搜索结果页面数据

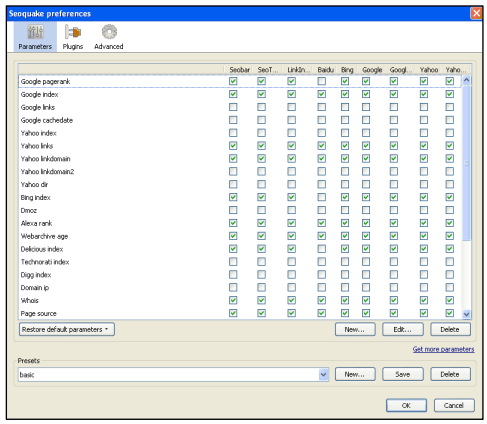


图 11-65　SeoQuake 插件设置

像前面提到的，这个工具查询数据量很大，经常造成用户 IP 地址被屏蔽。所以建议大家使用时，只查询自己想要看的数据，不要选择查询所有数据。

这个工具还有一个用途是，在雅虎链接查询工具中也可以使用，显示反向链接页面的 SEO 数据，同样包括 PR 值、链接数据等，这就可以非常方便地检查哪些链接的质量更高，如图 11-66 所示。

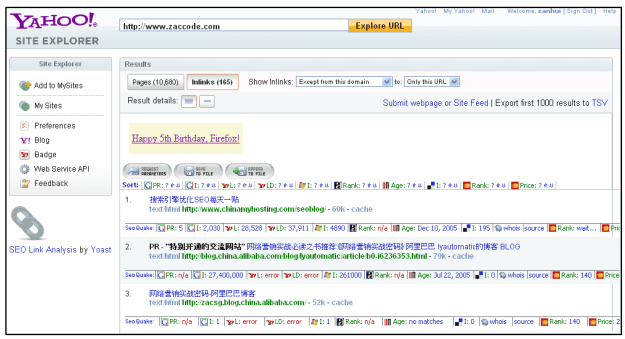


图 11-66　SeoQuake 插件用于 Yahoo! Site Explorer 页面

SeoQuake 插件功能非常强大，设置灵活，也是经常使用的 SEO 工具之一。

11.22　站长帮手

站长帮手提供不少与 SEO 有关的站长工具，如图 11-67 所示。网址是 http://www.linkhelper.cn/。

图 11-67　站长帮手

可以看到这是个典型的中国式网站，页面内容排得满满的，广告比较多，不注意看还真不知道工具在哪里。不过其中有几个工具比较有特色，值得介绍。

第一个是外链分析工具。输入一个网址后，这个工具查询外部链接，并显示这些外部链接的 PR 值、百度快照日期和锚文字，如图 11-68 所示。其中显示百度快照时间是前面介绍的其他工具所没有的。PR 值和锚文字当然也有重要参考价值。

图 11-68　站长帮手外链分析工具

缺点是锚文字信息需要人工单击“查看更多搜索引擎数据”链接才显示，而不是自动全部显示。

第二个有用的功能是网站收录查询，特点是在一个页面上，集成了几个中国主流搜索引擎的收录情况及反向链接数据，包括百度、Google、雅虎、搜狗、搜搜、有道和必应，还显示页面 PR 值、搜狗排名、Alexa 排名、百度快照日期，如图 11-69 所示。

图 11-69　站长帮手网站收录查询

像前面提到的，搜索引擎的 site:指令和 link:指令都不准确，尤其是查询反向链接的指令。但是所有这些数据放在一个页面，多少可以帮助站长了解网站大体情况。

第三个有用的是搜索引擎模拟抓取工具，也就是说显示出搜索引擎所看到的文字内容，剔除所有格式代码，如图 11-70 所示。

图 11-70　站长帮手模拟抓取

可惜工具只是把所有文字列出来，而没有标注哪些是 H 标签，哪些是黑体等。

11.23　关键词排名批量查询工具

批量查询某个网站的关键词排名是 SEO 人员必做的功课，不仅在研究竞争对手时要用到，自己网站优化前后都要定期查询，以监测优化效果，修改优化策略。这里介绍的批量查询工具就提供这样的功能，它是一款客户端软件。

下载地址是：http://www.flashplayer.cn/keywords/。

这个工具以前就叫关键字排名查询工具，2009 年 11 月份改名为观其站长工具箱。我猜想是这个工具被收购，或与其他 SEO 服务公司达成合作关系。可见制作 SEO 软件也是有前途的。

用户可以设置多个网站，每个网站设置多个关键词，然后查询在主流搜索引擎的排名，如图 11-71 所示。

图 11-71　关键词排名批量查询工具——观其站长工具箱

软件运行后，在界面内就可以查看这些关键词在各搜索引擎的排名情况，如图 11-72 所示。

图 11-72　观其站长工具箱查询结果

或者你也可以导出查询结果，生成一张更容易看清的排名列表，如图 11-73 所示。

序号	关键字/搜索引擎	百度	Google	Yahoo	soso	bing	搜狗	有道
1	seo	搜索排名为7 (推广1 - 自然6)	搜索排名为10	在1-10页中未查询到结果	在1-10页中未查询到结果	搜索排名为86	在1-10页中未查询到结果	搜索排名为9
2	网络营销	在1-10页中未查询到结果	搜索排名为40	搜索排名为63	在1-10页中未查询到结果	搜索排名为84	在1-10页中未查询到结果	在1-10页中未查询到结果
3	虚拟主机	搜索排名为73 (推广10 - 自然63)	搜索排名为40	数据获取失败	在1-10页中未查询到结果	在1-10页中未查询到结果	搜索排名为17	搜索排名为22
4	域名注册	搜索排名为62 (推广10 - 自然52)	搜索排名为64	无搜索结果或IP被屏蔽或其他错误	在1-10页中未查询到结果	在1-10页中未查询到结果	搜索排名为88	搜索排名为69
5	国外虚拟主机	搜索排名为12 (推广10 - 自然2)	搜索排名为6	数据获取失败	在1-10页中未查询到结果	搜索排名为7	搜索排名为1	搜索排名为1
6	搜索引擎优化	搜索排名为5 (推广3 - 自然2)	搜索排名为2	数据获取失败	搜索排名为74	搜索排名为3	在1-10页中未查询到结果	搜索排名为6

图 11-73　从观其站长工具箱导出查询结果

你还可以把数据输出为 Excel 格式文件。

除了偶尔数据获取失败这个小问题外，唯一比较大的缺憾是不能自动定期查询排名，并把历史排名数据存进一个数据库，供用户随时调用所有历史数据。用户需要人工输出数据，在 Excel 中合并数据。

11.24 SEOmoz 工具

外部链接是 SEO 最重要的因素之一，而链接数据却是 SEO 人员很难查询到的。

网站的内部结构、页面优化，我们都可以在网站本身看得很清楚。排名好的网站有哪些长处，只要仔细观察研究网站，就能学到很多东西。

到目前为止，链接因素还无法完全准确获得。一个特定网站在整个互联网的链接结构中到底处于什么样的位置，与哪些网站为邻，有多少个域名和页面链接过来，用的什么锚文字，这些数据只有搜索引擎自己知道。

雅虎链接工具算是比较准确的链接查询工具，但也有缺陷。第一，最多只能显示 1000 个链接。第二，显示的链接中包含了来自同一个域名的多个链接，也包含 nofollow 链接。所以对链接稍微多一点的网站来说，虽然能看到链接总数，却看不到超过 1000 个以后的具体链接来源。这 1000 个链接，来自同一个域名的可能就占了很大一部分，使得来自其他域名的链接显示不出来。

完整链接数据，搜索引擎大概永远也不会显示给我们看。要想获得准确的链接数据，从而把排名与链接之间的关系找出来，目前唯一可行的方法是自己建立页面及链接数据库。有几家公司做了这件事，SEOmoz 是其中之一。

SEOmoz 对所有 SEO 人员来说都不陌生，它是最有名的 SEO 服务商和博客之一。SEOmoz 创始人 Rand Fishkin 的帖子最有启发性，深受全世界 SEO 行业人员推崇，甚至连搜索引擎本身也乐于推荐。我在 2008 年厦门 SMX 大会上与 Google 美国总部中日韩文反垃圾组负责人朱建飞聊天时，他也提到，Google 算法有改动时，SEOmoz 往往能察觉到。这个评论令我印象深刻。

2007 年底，SEOmoz 为了扩展业务和转型，接受了 110 万美元的投资，开始大力研发 SEO 工具。2010 年 2 月，SEOmoz 更宣布不再提供 SEO 服务，完全转型为 SEO 工具提供商。

建立自己的链接索引库，就是 SEOmoz 工具的一个重点。它的链接数据库称为 Linkscape。Linkscape 于 2008 年 10 月推出，到 2011 年 11 月为止，收录了 1 亿 2 千 500 万个根域名的 380 亿个 URL，3870 亿个链接。虽然数据量与真正的搜索引擎相比还有一些差距，但大部分主要页面已经收录在它的数据库中。这些页面之间的链接关系在一定程度上反映了搜索引擎用来进行排名的链接关系。

mozRank 是 SEOmoz 依照 Google PR 的算法，但使用自己的链接数据计算出的页面重要性排名。一般来说 mozRank 值与 Google PR 值比较相近。如果 mozRank 值大大高于 Google PR 值，有可能是页面的反向链接中有一些是付费的或垃圾，被 Google 排除在

计算之外。Linkscape 工具可以在 2010 年 1 月推出的 Open Site Explorer 访问，网址是 http://www.opensiteexplorer.org/。完整版只有 SEOmoz 付费会员才能使用的。普通注册会员也可以使用几个简单的免费功能。

输入域名后，结果页面最上部显示页面基本信息，如图 11-74 所示。

图 11-74　Open Site Explorer 显示的页面链接概况

- Page Authority：基于链接计算的页面权重分数，满分为 100 分。
- Domain Authority：页面所在域名权重，满分为 100 分。
- Linking Root Domains：链接到这个页面 URL 的域名总数。
- Total Links：链接总数，包括 nofollow 的及 301 转向的。

Linking Root Domains 是连向这个页面的总域名数。前面提到过，链接的广泛性标准之一，就是链接过来的总域名数。有些域名可能提供很多链接，甚至一个大型网站加上一个全站链接，就有可能带来几百万外部链接。这样得来的几百万外部链接与来自几百万个不同域名的链接有天壤之别，很可能还不如几百个来自不同域名的链接。

我们都知道，链接就像是对页面的投票。来自同一个域名的再多链接，也还是同一个人或网站对你的投票。而来自几百个不同域名的链接，则意味着几百个不同的人对你的投票。

所以在某种程度上来说，反向链接绝对数量并不直接反映页面的真实链接广泛度，链接过来的总域名数才是更准确的数据。这个数据很少有工具能统计出来。

下面是 4 个选项卡，分别显示不同信息，如图 11-75 所示。

图 11-75　Open Site Explorer 显示的链接锚文字等信息

第一个选项卡，Linking Pages，显示这个页面的所有导入链接。链接页面信息包括页面标题、URL、锚文字、页面权重和域名权重。

三个下拉菜单可以分别选择：

- Followed + Nofollowed + 301，所有链接，包括 follow 的，nofollow 的，301 的。
- Only Nofollowed，只显示 nofollow 的链接。
- Only Followed，只显示 follow 的链接。
- Only 301，只显示 301 转向链接。

如图 11-76 所示，后两个菜单与 Yahoo Site Explorer 相似，用来选择：

- All Pages，来自所有页面的链接。
- External Pages Only，只显示外部链接。
- Internal Pages Only，只显示内部链接。

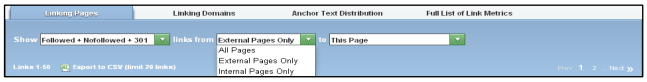


图 11-76　在 Open Site Explorer 选择链接来源

和：

- This Page，只显示连到这个页面的链接
- All Pages on the Subdomain，显示连到整个子域名所有页面的链接
- All Pages on the Root Domain，显示连到整个域名所有页面的链接

如图 11-77 所示。

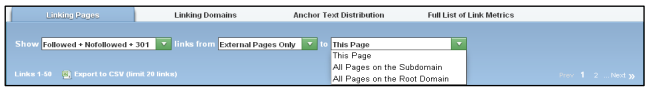


图 11-77　在 Open Site Explorer 选择页面或全站链接

第二个选项卡 Linking Domains，显示包含链接的所有域名。

第三个选项卡 Anchor Text Distribution，也很有价值，显示链接锚文字的分布，如图 11-78 所示。

第四个选项卡 Full List of Link Metrics，显示页面链接统计数字。

未登录访问者每天可以搜索 3 次，最多显示 200 个链接，其中 5 个链接显示详情（锚文字，页面权重，域名权重）。免费注册用户每天搜索次数无限制，最多显示 1000 个链接，其中 20 个链接显示详情。付费用户可以查询最多 1 万个链接，所有链接都显示详情。

Linking Pages | Linking Domains | Anchor Text Distribution | Full List of Link Metrics

Show anchor text Phrases for links to This Page

Top 100 Anchor Text Phrases Export to CSV (limit 20 anchors)

Anchor Text	Linking Root Domains Containing Anchor Text	Links Containing this Anchor Text
SEO每天一贴	142	2,914
Zac SEO每天一贴	141	308
搜索引擎优化SEO每天一贴	103	2,524
http://www.chinamyhosting.com/seoblog/	100	180
ZAC	78	173
Zac	75	451

图 11-78　Open Site Explorer 显示的链接锚文字分布

随着 Yahoo! Site Explorer 关闭，SEOmoz 的 Linkscape 和 Open Site Explorer 将成为最重要的 SEO 工具之一。

11.25　Backlink Watch

前面介绍了几个查询外部链接锚文字的工具，但使用上都有一点麻烦，需要用火狐浏览器加插件，或需要付费。Backlink Watch 这款反向链接工具使用就很简单，直接在网站上输入一个域名后，工具就返回所有外部链接列表，并同时显示锚文字、对方页面上的总导出链接数，以及链接是否带有 nofollow 属性，如图 11-79 所示。

网址是 http://www.backlinkwatch.com。

No.	Backlink URL	Anchor Text	PR	OBL	Flag
1	http://www.360seo.org/	××u		43	
2	http://www.lonmen.net/	点石互动		194	
3	http://robinliuchina.spaces.live.com/	点石互动		177	
4	http://www.inwayni.com/	点石大会SEO江湖		119	
5	http://www.sinzi.org/	点石		102	
6	http://www.23seo.org.cn/	点石互动		166	
7	http://www.seopt.cn/	点石互动		193	
8	http://blog.e4dai.com/	点石互动		246	
9	http://www.yijievalve.com/	××××SEO		51	
10	http://www.xiaofeng.org.cn/	点石博客		75	
11	http://www.qiyajun.cn/	点石互动		172	
12	http://spiderseo.com.cn/upload/%E5%A2%9E%E5%8A%A0%E5%8F.....				
13	http://www.chedong.com/blog/archives/001197.html	石头		139	nofollow
14	http://jiangzhanyong.com/2008/03/seo-tutorial-for-begin.....	点实互动		167	
15	http://sem.joyit.net/	点石互动		139	
16	http://www.phpcoding.cn/				
17	http://mysmth2003.spaces.live.com/			450	
18	http://www.zzdh5.cn/	××u××××		447	
19	http://5354.huychan.com/	点石互动		98	
20	http://www.etradewind.net/	点石互动		153	
21	http://www.newrain.net/blog/	点石		136	
22	http://www.shopex.cn/	××u××××		152	
23	http://www.seocool.com/	点石互动		103	
24	http://www.chinasearchdaily.com/	Dunsh Interactive		89	
25	http://googlechinawebmaster.com/	点石互动		282	

图 11-79　Backlink Watch 显示的外链锚文字

这个工具对中文网站支持有些问题，由于页面编码不同，有的页面锚文字显示是乱码，但大致上已经可以看出锚文字的构成，是简单研究外部链接锚文字的有效工具。

11.26 TouchGraph

做 SEO 的都知道应该尽量避免导出链接到坏邻居，也就是说，你的网站经常链接到作弊网站，你的网站也会受牵连。但我们很难确切地知道一个网站到底和哪些其他网站是邻居。一个网站上的导入导出链接数量可能十分巨大，我们很难掌握链接的整体情况。

TouchGraph 这个工具用图像方式非常直观地显示出一个网站是和哪些其他网站处在同一个社区，链接邻居是哪些，这些网站之间的链接关系怎样。

网址是 http://www.touchgraph.com/TGGoogleBrowser.html。

如图 11-80 所示，《网络营销实战密码》读者交流网站 zaccode.com 周围都是一些与 SEO、网络营销等互联网有关的邻居，还算比较健康。

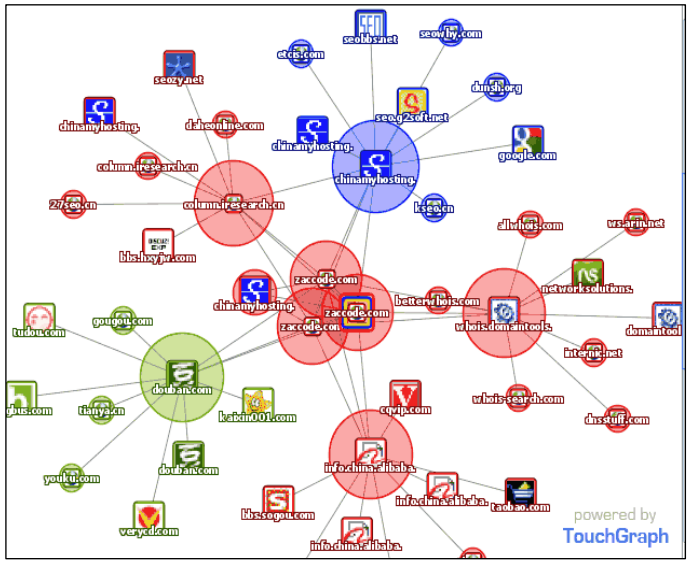


图 11-80 TouchGraph 显示的网站链接关系

11.27 Quintura

前面提到过语义分析与网页文案写作的问题，Quintura 这款工具就可以用来显示与查询词有语义关系的更多关键词。比如输入“网络营销”，工具返回如图 11-81 所示的这些语义相关词汇。

请注意，这与通常关键词扩展工具中得到的数据并不一样。在关键词扩展工具中输入“网络营销”一般会扩展出“网络营销工具”、“北京网络营销”等相关搜索词。而 Quintura 工具返回的是语义相关词，诸如“维基百科”、“成功案例”，很难在关键词工具中得到，站长自己也比较难想到这些词之间是有语义关系的。

图 11-81　Quintura 生成语义相关词

这个工具是从大量数据中分析出语义关系，而不是靠人脑判断有没有关系。掌握了这些有语义关系的词汇，有助于文案写作，使主题更为集中。

网址是 http://www.quintura.com/。

11.28　Google Ad Planner

Google Ad Planner 本来是 Google 提供给广告商寻找合适的广告发布网站的工具。要寻找合适的广告网站，就自然要显示对方网站的一些数据，这就刚好可以被 SEO 人员用来做竞争对手调查，如图 11-82 所示。

网址是 https://www.google.com/adplanner/。

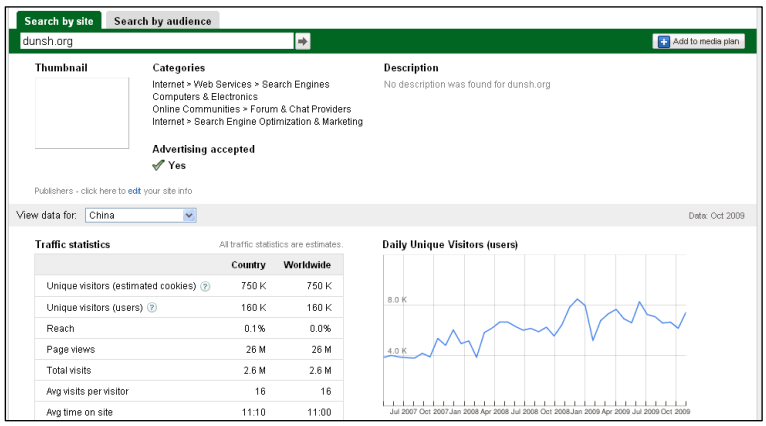


图 11-82　Google Ad Planner 显示的网站流量数据

Google Ad Planner 可以显示一个网站的流量数据，与 Google 趋势类似。同时还显示页面访问总数，每个用户访问的页面数，以及在网站上停留的时间。Google Ad Planner

也给网站进行了大致归类，如点石互动网站被归类为搜索引擎优化和推广类别，但有时候这个类别也分得很不靠谱儿。

11.29 Majestic SEO

Majestic SEO 是一款深度挖掘网站外部链接数据的工具，虽然详细数据需要付费，不过价格不高，很多情况下物超所值。

网址是 http://www.majesticseo.com。

Majestic SEO 可以显示网站的链接数据，既可以查询自己网站外链，也可以查询别人网站数据。如果是自己的网站，可以免费获得链接数据，只要在服务器上放一个验证文件就可以了。

要查询别人网站需要付费。2010 年以前是根据链接数目收费，比如 SEO 每天一贴链接少，只需要两英镑。

如果想查询百度这种反向链接达到上亿级别的网站，报价是 400 英镑，如图 11-83 所示。

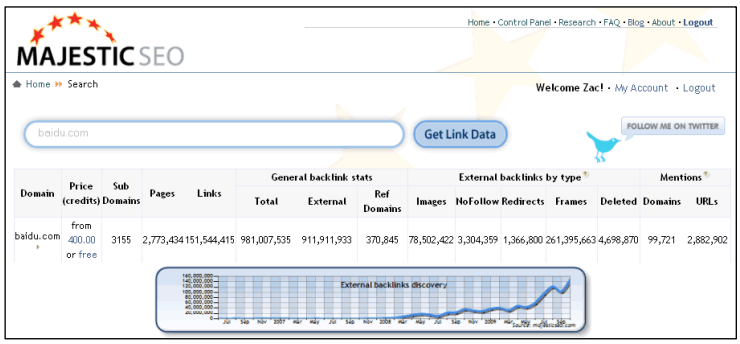


图 11-83 Majestic SEO

58 同城这种千万以上链接级别的网站，报价为 50 英镑。站长站链接 200 多万，报价是 35 英镑。链接 200 万以下的网站，通常就只需要付一两英镑而已。

2010 年后 Majestic SEO 收费方法改为会员制，按月付费。银卡会员每个月可以查询获得 60 份链接详细报告，月费 29.99 英镑。金卡会员每月 300 份报告，99.99 英镑。白金卡会员 950 份报告，250 英镑。

虽然使用价格没有以前划算，但对较大型网站来说，几十万、上百万链接相当难得，要建这么多链接所花费的时间、人力成本相当高。花 30 英镑获得几十个竞争对手的相对准确的链接数据，找出竞争对手为什么排名好，从哪些地方获得链接，自己又能从这些链接源获得更多链接，显然物超所值。

任何没有登录的访问者都可以查询任何网站的外部链接基本数据，如总链接数、总外部链接数、带来链接的总域名数等，但看不到链接列表和更详细的情况。

Majestic SEO 公司宣称本身并不是 SEO 工具提供商，它本身就是个搜索引擎。目前数据库中有 11000 千亿网址，这个数字离真正的搜索引擎也差不多了，不过我们也无法验证。

验证了自己的网站或付费获得其他网站数据访问权后，用户可以在工具后台看到很多详细数据，左上角有菜单可以方便切换，如图 11-84 所示。

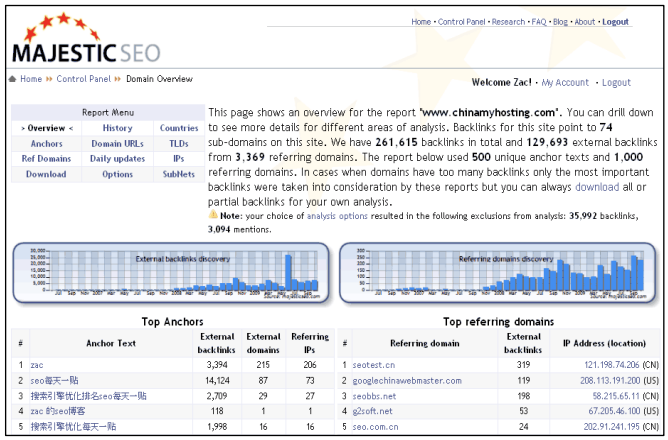


图 11-84　Majestic SEO 后台界面

下面简单介绍主要数据。

链接增长历史（Backlinks History），用图表的方式非常直观地显示出网站外部链接的增长（或下降）趋势，如图 11-85 所示。

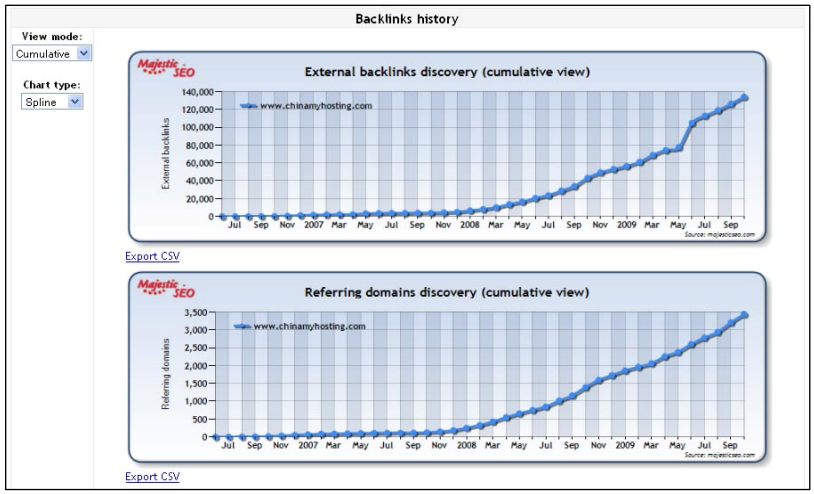
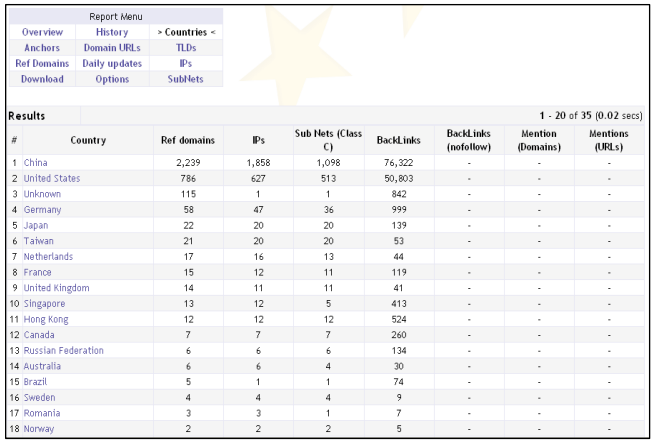


图 11-85　外链增长趋势

前面说过，随时间的变化是健康外部链接构成因素之一。一个好的网站，通常外部链接随时间逐渐积累，越来越多，较少跳跃式增长，链接减少就更少见了。

链接历史也能反映出自己网站链接建设效果如何，竞争对手的哪些营销活动带来了大量外部链接，链接诱饵策划和执行得怎么样，链接数量剧烈变化与排名大幅波动是否有时间关系，等等。

链接地理分布。单击 Countries 菜单可以看到来自不同国家的链接数目，地理分布是根据链接源域名服务器所在位置区分的，如图 11-86 所示。



Report Menu

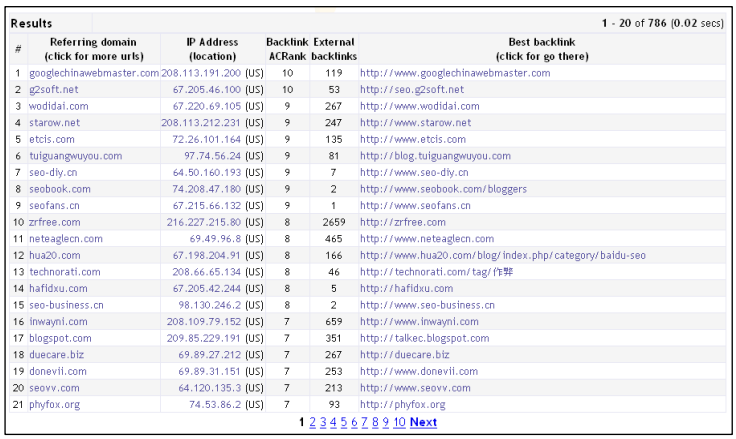
Overview	History	> Countries <
Anchors	Domain URLs	TLDs
Ref Domains	Daily updates	IPs
Download	Options	SubNets

Results 1 - 20 of 35 (0.02 secs)

#	Country	Ref domains	IPs	Sub Nets (Class C)	BackLinks	BackLinks (nofollow)	Mention (Domains)	Mentions (URLs)
1	China	2,239	1,858	1,098	76,322	-	-	-
2	United States	786	627	513	50,803	-	-	-
3	Unknown	115	1	1	842	-	-	-
4	Germany	58	47	36	999	-	-	-
5	Japan	22	20	20	139	-	-	-
6	Taiwan	21	20	20	53	-	-	-
7	Netherlands	17	16	13	44	-	-	-
8	France	15	12	11	119	-	-	-
9	United Kingdom	14	11	11	41	-	-	-
10	Singapore	13	12	5	413	-	-	-
11	Hong Kong	12	12	12	524	-	-	-
12	Canada	7	7	7	260	-	-	-
13	Russian Federation	6	6	6	134	-	-	-
14	Australia	6	6	4	30	-	-	-
15	Brazil	5	1	1	74	-	-	-
16	Sweden	4	4	4	9	-	-	-
17	Romania	3	3	1	7	-	-	-
18	Norway	2	2	2	5	-	-	-

图 11-86　外链来自哪些国家

笔者博客 SEO 每天一贴，很显然来自中国的链接最多。单击来自美国的链接，可以看到服务器在美国的哪些网站链接到我的博客，如图 11-87 所示。



Results 1 - 20 of 786 (0.02 secs)

#	Referring domain (click for more urls)	IP Address (location)	Backlink ACRank	External backlinks	Best backlink (click for go there)
1	googlechinawebmaster.com	208.113.191.200 (US)	10	119	http://www.googlechinawebmaster.com
2	g2soft.net	67.205.46.100 (US)	10	53	http://seo.g2soft.net
3	wodidai.com	67.220.69.105 (US)	9	267	http://www.wodidai.com
4	starow.net	208.113.212.231 (US)	9	247	http://www.starow.net
5	etcis.com	72.26.101.164 (US)	9	135	http://www.etcis.com
6	tuiguangwuyou.com	97.74.56.24 (US)	9	81	http://blog.tuiguangwuyou.com
7	seo-diy.cn	64.50.160.193 (US)	9	7	http://www.seo-diy.cn
8	seobook.com	74.208.47.180 (US)	9	2	http://www.seobook.com/bloggers
9	seofans.cn	67.215.66.132 (US)	9	1	http://www.seofans.cn
10	zrfree.com	216.227.215.80 (US)	8	2659	http://zrfree.com
11	neteaglecn.com	69.49.96.8 (US)	8	465	http://www.neteaglecn.com
12	hua20.com	67.198.204.91 (US)	8	166	http://www.hua20.com/blog/index.php/category/baidu-seo
13	technorati.com	208.66.65.134 (US)	8	46	http://technorati.com/tag/作弊
14	hafidxu.com	67.205.42.244 (US)	8	5	http://hafidxu.com
15	seo-business.cn	98.130.246.2 (US)	8	2	http://www.seo-business.cn
16	inwayni.com	208.109.79.152 (US)	7	659	http://www.inwayni.com
17	blogspot.com	209.85.229.191 (US)	7	351	http://talkec.blogspot.com
18	duecare.biz	69.89.27.212 (US)	7	267	http://duecare.biz
19	donevii.com	69.89.31.151 (US)	7	253	http://www.donevii.com
20	seovv.com	64.120.135.3 (US)	7	213	http://www.seovv.com
21	phyfox.org	74.53.86.2 (US)	7	93	http://phyfox.org

1 2 3 4 5 6 7 8 9 10 Next

图 11-87　来自美国的外链列表

排在第一的是权重最高的 Google 网站管理员博客。从列表中还可以看到，原来不少国内的 SEO 其实把博客或网站放在了美国。

单击排在第二的 DavidYin 的博客（g2soft.net），可以看到 DavidYin 博客上哪些页面有 SEO 每天一贴的链接，链接到哪个 URL，锚文字是什么，如图 11-88 所示。

Results 1 - 20 of 43 (0.05 secs)

#	Source URL	Source ACRank	Anchor text	Date Found	Target URL	Target ACRank	Source Flags Redirect	Frame	NoFollow	Images	Deleted	AltText
1	http://seo.g2soft.net	10	zac的每天一贴	9 Oct 2009	http://www.chinamyhosting.com/seoblog	11						
2	http://seo.g2soft.net/baidu	6	blog	7 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	7						
3	http://seo.g2soft.net/baidu	6	搜索引擎优化排名seo每天一贴	7 Sep 2009	http://www.chinamyhosting.com/seoblog	11						
4	http://seo.g2soft.net/2006/09/04/blog-op...	4	blog优化有很大的局限	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	5						
5	http://seo.g2soft.net/2006/10/27/zac-blo...	4	搜索引擎优化排名seo每天一贴	6 Sep 2009	http://www.chinamyhosting.com/seoblog	11						
6	http://seo.g2soft.net/2006/10/27/zac-blo...	4	blog	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	7						
7	http://seo.g2soft.net/2008/01/02/chinamy...	4	正式上线了	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	5						
8	http://seo.g2soft.net/2008/03/01/write-b...	4	zac	6 Sep 2009	http://www.chinamyhosting.com/seoblog	11						

图 11-88　来自特定网站的链接详情

Majestic SEO 工具的数据比较全，经常可以看到不少自己无法察觉的外部链接来源。比如 SEO 每天一贴竟然还有来自巴西、瑞典的链接。

链接锚文字（Anchor），使用每种锚文字的外部链接总数（External backlinks），使用每种锚文字的域名总数（Referring domains）等，如图 11-89 所示。

Results 1 - 20 of 500 (0.02 secs)

#	Anchor Text	Check	External backlinks	Referring domains	Referring IPs	Backlinks ACRank spread 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	zac	(G | Y)	3394	215	206	799 (+1,956)	70 (+256)		140 (+18)	63	44	25	18	2	2	1					
2	seo每天一贴	(G | Y)	14124	87	73	16 (+12,337)	2 (+1,317)		229 (+82)	82	23	9	15	6	5	1					
3	搜索引擎优化排名seo每天一贴	(G | Y)	2709	29	27	37 (+2,410)	4 (+143)		74		21	11	6	2		1					
4	zac 的seo博客	(G | Y)	118	1	1	(+2)	(+1)		6		17	31	21	37	2		1				
5	搜索引擎优化每天一贴	(G | Y)	1998	16	16	11 (+1,601)	129 (+77)		140		21	8	7	2		2					
6	搜索引擎优化	(G | Y)	2201	18	16	17 (+2,084)	(+71)		19 (+2)		3	1	2			2					
7	搜索引擎优化seo每天一贴	(G | Y)	8679	114	104	147 (+7,436)	10 (+642)		200 (+88)		111	14	11	17	2	1					
8	搜索引擎优化seo每日一帖	(G | Y)	1595	5	5	12 (+1,341)	7 (+138)		68		15	6	5	2		1					
9	zac-每天一贴	(G | Y)	226	1	1	(+113)	8		44		60					1					
10	国外虚拟主机域名注册	(G | Y)	218	18	15	1 (+111)	49		22		17	7	4	6		1					
11	seo每日一帖	(G | Y)	882	13	8	(+808)	(+46)		16		2	3	1	5		1					
12	zac@seo每天一帖	(G | Y)	145	7	4	8 (+94)	(+33)		7		2					1					
13	zac博客	(G | Y)	729	7	7	2 (+619)	(+34)		5 (+4)		47	8	3	6	1						
14	中新网络科技	(G | Y)	221	14	14	14 (+171)	6		6		2		12	9	1						
15	seoblog	(G | Y)	262	1	1	(+185)	(+44)		27		5				1						
16	http://www.chinamyhosting.com/seobl...	(G | Y)	310	11	11	19 (+261)	(+19)		8		2				1						
17	zac's blog	(G | Y)	191	1	1	(+130)	(+34)		20		6				1						
18	zac的博客	(G | Y)	413	18	18	9 (+294)	1 (+77)		17 (+1)		7	2	1	4							

图 11-89　外链锚文字分布列表

从图中可以看到，链接到我博客的锚文字最多的是“SEO 每天一贴”，共有 14000 多个链接，但以“SEO 每天一贴”作为锚文字的链接只来自 87 个域名。使用“zac”做锚文字的链接总数少很多，只有 3000 多，但域名总数达 200 多。下面还有更多形式的锚文字，五花八门。锚文字越呈现多样性，链接越显得自然。

你也可以点击这些锚文字，进一步查看使用某个特定锚文字的链接都有哪些。如图 11-90 所显示的是使用“URL 网址规范化”作为锚文字的链接。

Report Menu		
Overview	History	Countries
Anchors	Domain URLs	TLDs
Ref Domains	Daily updates	IPs
Download	Options	SubNets

Below you can see sorted by ACRank site backlinks pointing to URLs contained in the report '**www.chinamyhosting.com**' with anchor text of '**url网址规范化**'. Domain clustering is **ON** (turn it OFF), multiple urls from the same domain will not be shown, only the highest ACRanked one will be shown.

Results 1 - 16 of 16 (0.06 secs)

#	Source URL	Source ACRank	Date Found	Target URL	Target ACRank	Source Flags: Redirect	Frame	NoFollow	Images	Deleted	AltText
1	http://www.osxcn.com/wordpress/let-wordp... (+4 filtered)	7	7 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	7						
2	http://www.ulland.com/wordpress/seo-in-w... (+539 filtered)	4	19 Jun 2009	http://www.chinamyhosting.com/seoblog/20...	7						
3	http://www.xqxp.com/blog/?p=320	4	20 Nov 2007	http://www.chinamyhosting.com/seoblog/20...	7						
4	http://blog.designlinks.cn/wordpress/let... (+1 filtered)	3	30 May 2008	http://www.chinamyhosting.com/seoblog/20...	7						
5	http://lilhai.com/seo_from_today_permali... (+1 filtered)	3	17 Mar 2008	http://www.chinamyhosting.com/seoblog/20...	7						
6	http://www.cblog.name/archives/wordpress...	3	25 Apr 2009	http://www.chinamyhosting.com/seoblog/20...	7						
7	http://cocolike.com/blog/seo-url-canonic... (+1 filtered)	1	7 Feb 2008	http://www.chinamyhosting.com/seoblog/20...	7						
8	http://www.ankr.cn/feed (+1 filtered)	0	12 Jul 2009	http://www.chinamyhosting.com/seoblog/20...	7						
9	http://orrio.cn/3-plugins-for-wordpress-...	0	30 Jan 2008	http://www.chinamyhosting.com/seoblog/20...	7						
10	http://gouwukuang.cqmmgo.com/2008/05/301...	0	6 Jan 2009	http://www.chinamyhosting.com/seoblog/20...	7						

图 11-90　使用特定锚文字的外链详情

所有链接列表（Domain URLs），显示网站所有获得了外部链接的网址，以及获得链接的情况。如图 11-91 所示，链接到“SEO 每天一贴”博客首页的最多，共有来自 1992 个域名的 54757 个链接。你还可以单击每一个 URL，查看这个 URL 的具体链接情况。如图 11-92 所示，为“SEO 每天一贴”博客首页的外部链接列表，包含来源网址和锚文字。

Results 1 - 20 of 500 (0.02 secs)

#	URL	ACRank	External backlinks	Referring domains	Referring IPs	Backlinks ACRank spread: 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	http://www.chinamyhosting.com/seoblog	11	54757	1992	1,628	(+47,693)	(+4,509)		1,427 (+219)	458	177	98	132	17	21	6					
2	http://www.chinamyhosting.com	9	5501	707	638	856 (+3,591)	561		283	124	28	25	23	6	4						
3	http://www.chinamyhosting.com/seoblog/2006/07/27/site-struct...	8	309	176	169	279	11		8	4	4	1	2								
4	http://www.chinamyhosting.com/seoblog/category/marketing	8	619	177	165	574	29		12	1		3									
5	http://www.chinamyhosting.com/seoblog/2006/04/08/google-pr	8	462	135	127	423	24		8	3	3	1									
6	http://www.chinamyhosting.com/seoblog/2006/04/11/quality-inc...	8	322	155	148	306	7		3	3	3										
7	http://www.chinamyhosting.com/seoblog/about	8	3022	148	140	1,542 (+1,226)	199		50	3	2										
8	http://www.chinamyhosting.com/seoblog/2006/04/12/301-redirec...	8	311	120	114	267	30		9	3	2										
9	http://www.chinamyhosting.com/seoblog/2006/04/24/link-exchan...	8	273	152	142	263	6		2	1	1										
10	http://www.chinamyhosting.com/seoblog/2008/03/14/blog-seo	7	153	40	39	131	11		4	2	4			1							
11	http://www.chinamyhosting.com/seoblog/2006/06/23/google-know...	7	98	61	60	79	3		5	2	2	4	3								

图 11-91　所有外链列表

Results 1 - 20 of 1,992 (0.06 secs)

#	Source URL	Source ACRank	Anchor Text	Date Found	Source Flags: Redirect	Frame	NoFollow	Images	Deleted	AltText
1	http://www.zzdh.net (+9 filtered)	10	seo每天一贴	21 Oct 2009						
2	http://www.seobbs.net (+8 filtered)	10	[seo每天一帖]	9 Sep 2009						
3	http://seo.g2soft.net (+7 filtered)	10	zac的每天一贴	9 Oct 2009						
4	http://www.googlechinawebmaster.com (+114 filtered)	10	zac 的seo博客	20 Sep 2009						
5	http://www.seotest.cn (+5 filtered)	10	搜索引擎优化排名seo每天一贴	9 Oct 2009						
6	http://www.seo.com.cn (+3 filtered)	10	zac	12 Sep 2009						
7	http://www.piaoyi.org (+12 filtered)	9	seo每天一贴	15 Sep 2009						
8	http://www.360seo.org	9	zac	7 Sep 2009						
9	http://www.zaiqiang.com	9	搜索引擎优化	29 Sep 2009						
10	http://www.wodidai.com (+50 filtered)	9	seo每天一贴	2 Oct 2009						
11	http://blog.tuiguangwuyou.com	9	seo每天一贴	20 Sep 2009						
12	http://www.timev.com	9	zac营销博客	20 Sep 2009						
13	http://www.ktwo888.com	9	搜索引擎优化	10 Sep 2009						
14	http://seo.highdiy.com (+10 filtered)	9	搜索引擎优化每天一贴	7 Sep 2009						
15	http://www.guanzhipeng.com (+4 filtered)	9	搜索引擎优化seo每天一贴	21 Sep 2009						
16	http://www.etcis.com (+71 filtered)	9	搜索引擎优化每天一贴	18 Sep 2009						
17	http://www.bloghuman.com (+96 filtered)	9	搜索引擎优化seo每日一帖	19 Oct 2009						
18	http://www.100ye.com/shopdetail/index/in...	9	zac的seo优化	4 Sep 2008						
19	http://www.seovip.cn/blog	9	zac每天一帖	13 Sep 2009						
20	http://www.seopm.cn/site/Homepage-optimi...	9	seo每天一贴	16 Jun 2009						

1 2 3 4 5 6 7 8 9 10 **Next**

图 11-92　特定页面的外链详情

链接来源域名种类（TLDs），可以查看所有外部链接来自哪种类型的域名，如图 11-93 所示。

#	Short TLD	Long TLD	Referring domains	External backlinks	Backlinks ACRank spread 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	asia	.asia	1	1				1												
2	at	.at	1	1	1															
3	au	.com.au	1	17	17															
4	be	.be	2	4	4															
5	bg	.bg	1	1	1															
6	biz	.biz	5	390	216	117		25	31			1								
7	br	.com.br	2	32	32															
8	ca	.ca	3	4	2	2														
9	cc	.cc	19	136	130	2		2	2											
10	cc	.co.cc	3	4	3				1											
11	ch	.ch	1	2	2															
12	cn	.cn	795	51,926	49,494	1,507		648	159	56	25	22	6	8	1					
13	cn	.com.cn	169	5,614	5,303	168		94	23	14	3	6	1	1	1					
14	cn	.edu.cn	3	32	32															
15	cn	.net.cn	33	211	192	2		12	1	2		1		1						
16	cn	.org.cn	40	1,258	1,054	136		47	12	2	2	5								
17	com	.com	1,628	45,935	40,375	3,566		1,203	385	171	98	107	16	13	1					
18	de	.de	16	44	42	2														
19	dk	.dk	3	9	9															
20	ee	.ee	1	1	1															
21	eg	.com.eg	1	1	1															
22	es	.es	1	1	1															
23	eu	.eu	5	8	7			1												
24	fm	.fm	1	14	14															
25	fr	.fr	2	6	4			1	1											
26	gr	.gr	1	2	2															

图 11-93　外链来源域名种类分布

链接来源域名列表（Referring Domains），这一项显示所有链接到你网站的域名，如图 11-94 所示。

#	Referring domain (click for more urls)	IP Address (location)	Backlink ACRank	External backlinks	Best backlink (click for go there)	Backlinks ACRank spread 0	1	2	3	4	5	6	7	8	9	10	11
1	seotest.cn	121.198.74.206 (CN)	10	319	http://www.seotest.cn	225 (+55)	7 (+3)		2 (+1)	9 (+4)	5 (+2)	3 (+2)				1	
2	googlechinawebmaster.com	208.113.191.200 (US)	10	119	http://www.googlechinawebmaster.com		1 (+1)		6	18	31	21 (+1)	37	2		1	
3	seobbs.net	58.215.65.11 (CN)	10	198	http://www.seobbs.net	32 (+154)	1 (+2)		7				1			1	
4	g2soft.net	67.205.46.100 (US)	10	53	http://seo.g2soft.net	8 (+6)	7 (+6)		12 (+4)	6 (+1)		1 (+1)				1	
5	seo.com.cn	202.91.241.195 (CN)	10	24	http://www.seo.com.cn	6 (+9)	1		2 (+1)	2	2					1	
6	zzdh.net	113.11.200.103 (CN)	10	30	http://www.zzdh.net	(+19)	(+1)		2	2	3		2			1	
7	starow.net	208.113.212.231 (US)	9	247	http://www.starow.net	4 (+123)	8		44 (+1)	61 (+3)		1		1 (+1)			
8	timev.com	121.10.121.31 (CN)	9	392	http://www.timev.com	108 (+264)	5 (+12)		2					1			
9	bloghuman.com	211.136.108.135 (CN)	9	1678	http://www.bloghuman.com	47 (+1,378)	5 (+145)		68 (+3)	15	8	6	2	1			
10	seovip.cn	222.76.217.169 (CN)	9	95	http://www.seovip.cn/blog	92 (+2)								1			
11	etcis.com	72.26.101.164 (US)	9	135	http://www.etcis.com	(+42)	(+21)		56	5	4	6		1			
12	lesishu.cn	58.215.65.11 (CN)	9	102	http://www.lesishu.cn	8 (+27)	3 (+12)		17 (+14)	8 (+8)	1 (+1)	1 (+1)		1			
13	wodidai.com	67.220.69.105 (US)	9	267	http://www.wodidai.com	(+197)	(+19)		42	8				1			
14	sinatea.com	115.28.150.81 (CN)	9	29	http://www.sinatea.com	(+2)	24		2					1			
15	highdiy.com	218.57.8.39 (CN)	9	114	http://seo.highdiy.com	2 (+89)	1 (+2)		8	3	2 (+1)	3 (+1)	1	1			
16	kseo.cn	222.186.191.112 (CN)	9	25	http://www.kseo.cn	(+2)	(+1)		1 (+1)	3	4 (+10)	2		1			
17	plaoyi.org	202.75.217.53 (CN)	9	22	http://www.plaoyi.org	1 (+4)	1 (+1)		10	1	2	1		1			
18	ipggo.cn	125.64.92.26 (CN)	9	30	http://www.ipggo.cn	(+17)	(+5)		5	1	1			1			
19	guanzhipeng.com	221.194.44.143 (CN)	9	564	http://www.guanzhipeng.com	(+505)	(+54)		4					1			
20	seo-diy.cn	64.50.160.193 (US)	9	7	http://www.seo-diy.cn	3 (+2)	(+1)							1			

1 2 3 4 5 6 7 8 9 10 **Next**

图 11-94　链接来源域名列表

从图中可以看到，给“SEO 每天一贴”带来链接比较多，权重又比较高，排在第一的是 seotest.cn。如果通常博客的 blogroll 中出现你的链接，一个网站就可以带来很多链接。单击 seotest.cn 可以查看所有来自 seotest.cn 网站的链接情况，包括链接源网址、锚文字及链接指向的网址，如图 11-95 所示。

每天更新（Daily updates）。Majestic SEO 的数据库也是每天更新的，在 Daily updates 部分可以查看每一天你的网站又增加了哪些新链接。

需要指出的是，这里所说的每天更新，指的是 Majestic SEO 新发现的链接，其中既包括真的是刚刚出现的链接，也包括早就存在，但 Majestic SEO 刚刚检测收录到的链接。图 11-96 所示是 SEO 每天一贴在 2009 年 11 月 16 日被发现的新链接。

选项（Options），如图 11-97 所示。

#	Source URL	Source ACRank	Anchor text	Date Found	Target URL	Target ACRank	Source Flags Redirect	Frame	NoFollow	Images	Deleted	AltText
1	http://www.seotest.cn	10	搜索引擎优化排名seo每天一贴	9 Oct 2009	http://www.chinamyhosting.com/seoblog	11						
2	http://feed.seotest.cn	6	到底seo是什么	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	6						
3	http://feed.seotest.cn	6	zac的"搜索引擎每日一贴"	6 Sep 2009	http://www.chinamyhosting.com/seoblog	11						
4	http://www.seotest.cn/blog/zac-keyword-b...	6	关键词布局	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	6						
5	http://www.seotest.cn/rss.xml	6	到底seo是什么	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	6						
6	http://www.seotest.cn/rss.xml	6	zac的"搜索引擎每日一贴"	6 Sep 2009	http://www.chinamyhosting.com/seoblog	11						
7	http://www.seotest.cn/blog/301topr.html	5	zac的话	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	7						
8	http://www.seotest.cn/blog/fencl-xiezuo....	5	到底seo是什么	10 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	6						
9	http://www.seotest.cn/blog/fencl-xiezuo....	5	zac的"搜索引擎每日一贴"	10 Sep 2009	http://www.chinamyhosting.com/seoblog	11						
10	http://www.seotest.cn/blog/seo-fuwu-qian....	5	seo服务前景不妙	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	6						
11	http://www.seotest.cn/catalog.asp?cate=1	5	"nofollow"更详细说明	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	7						
12	http://www.seotest.cn/catalog.asp?cate=1	5	zac的话	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	7						
13	http://www.seotest.cn/catalog.asp?tags=关...	5	关键词布局	6 Sep 2009	http://www.chinamyhosting.com/seoblog/20...	6						

图 11-95　来自特定域名的链接详情

Report Menu

Overview	History	Countries
Anchors	Domain URLs	TLDs
Ref Domains	Daily updates	IPs
Download	Options	SubNets

Below you can see sorted fresh backlinks pointing to DD2F1B267CD533E4 found on **2009-11-16**. By switching off domain clustering you will be able to see target urls for each of the source urls that were found.

Domain clustering is **ON** (turn it OFF), multiple urls from the same domain will not be shown, only the highest ACRanked one will be shown.

Results　　1 - 20 of 64 (0.02 secs)

#	Source URL	Source ACRank	Anchor Text	Date Found	Source Flags Redirect	Frame	NoFollow	Images	Deleted	AltText
1	http://www.123tz.cn/bbs/read.php?tid=15057	0	www.chinamyhosting.com/seoblog	16 Nov 2009						
2	http://bbs.179179.cn/viewthread.php?action=printable&tid=170...	0		16 Nov 2009						✓
3	http://21cdata.cn/news/view.asp?id=225	0	http://www.chinamyhosting.com/...	16 Nov 2009						
4	http://www.adminkc.cn/viewthread.php?tid=34720	0		16 Nov 2009						✓
5	http://aliqidian.cn	0	seo每天一贴	16 Nov 2009						
6	http://www.16seo.com.cn/?p=107 (+37 filtered)	0	seo每天一贴	16 Nov 2009						
7	http://so.chinawj.com.cn/google.aspx?wd=PPC	0	ppc竞价排名搜索引擎– 搜索引擎...	16 Nov 2009						
8	http://www.gdecs.com.cn/asblog/tag/blockip	0		16 Nov 2009						✓
9	http://itbear.com.cn/Archiver/2007-10/5764.html	0	google pr值越来越不重要了	16 Nov 2009						

图 11-96　新增加的外链

Home » Control Panel » Report analysis options　　Welcome Zac! · My Account · Logout

Report Menu

Overview	History	Countries
Anchors	Domain URLs	TLDs
Ref Domains	Daily updates	IPs
Download	> Options <	SubNets

You can customise domain analysis for the report "**www.chinamyhosting.com**" in order to focus on the elements of primary interest. Once you have made changes to the options it is necessary to force domain analysis for the reports to be updated.

Target URL options

Limit analysis to these target URLs:

Note: all urls must belong to the domain that is being analysed. If none specified then all domain urls will be analysed. Maximum of 100 urls allowed. You can use subdomain without http:// to limit analysis to all urls on that subdomain, ie: **www.example.com** will limit analysis to all target urls on this subdomain only, however http://www.example.com will limit analysis to only that URL.

Limit analysis to homepages only:

☐ If checked then only homepages of the domain and any subdomains will be considered for this analysis. This option is ideal if you want to create a list of all subdomains on this domain.

Filtering rules for the Target URLs that contain specified text:

Include, if URL contains:　　**Exclude, if URL contains:**

Note: you can use commas to provide multiple match strings. If you prefix URLs with ^ then exact match will be required, ie: ^**http://www.example.com** will attempt to match that exact URL.

Source URL options

图 11-97　Majestic SEO 选项设置

在这里你可以设置显示哪些链接，通过过滤器可以包含或排除某些链接。因为 Majestic SEO 提供的数据十分庞大，通过不同过滤可以比较方便地查看你比较关注的数据。

使用 Majestic SEO 工具获得竞争对手外部链接资料，可以极大地扩展你自己获得外部链接的来源及可能性。Majestic SEO 提供的数据确实是很少能在其他地方找到的。

11.30 追词

追词是很不错的一个集成 SEO 工具网站。VIP 用户需要付费，但大部分功能免费用户也可以使用，如图 11-98 所示。

网址是 http://www.zhuici.com/。

图 11-98 追词网控制中心

所有界面都是中文的，清晰易懂，就不一一介绍了，只简单介绍几个有意思的功能。

作为国内站长开发的工具，提供中文搜索引擎数据当然是特点之一。比如关键词挖掘部分，显示百度及 Google 的收录数，广告数目，排在第一的页面及反向链接数。不仅显示 Google PR 值，还显示搜狗排名值，如图 11-99 所示。

图 11-99 追词网关键词挖掘功能

关键词部分另一个有意思的功能是智能报价。用户输入关键词工具，追词根据自己的计算公式报给你一个要优化这个词的价格，如图 11-100 所示。

图 11-100　追词网智能报价功能

当然，这个价格仅供参考，不同地区、不同公司提供 SEO 服务的价格相差很大。

大部分免费工具都不提供数据存储，追词网则提供这个很重要的功能。监测项目把用户感兴趣的数据存在服务器上，供以后随时调阅，比如图 11-101 所示的反向链接变化情况。

图 11-101　追词网反链监测

排行榜部分有两个有意思的功能。一个功能是智能排名检测，可以查询一个网站在百度获得的比较好排名，如图 11-102 所示。

图 11-102　追词网智能排名检测

用户输入一个域名，工具将返回这个域名在百度都有哪些关键词排名不错。很多时候，SEO 人员需要知道竞争对手或自己网站整体上获得了哪些关键词排名。自己的网站还可以查看流量和日志，竞争对手网站只能从搜索结果页面查看。一个一个地人工查询显然工作量太大，而且不容易全面，有很多关键词排名是我们想不到的。追词的这个功能迅速告诉你特定网站已经有哪些排名。

另外一个功能是优质关键词。用户输入一个搜索次数范围，追词将显示符合这个搜索次数范围的关键词，和 Google 热榜、百度排行榜等类似，如图 11-103 所示。

关键词	搜索次数	百度价格	Google价格	百度收录数	谷歌收录数
鼠标	2862	0.61	N/A	58400000	37500000
黑塔利亚	3127	N/A	N/A	370000	6250000
魅力云江紫	5443	N/A	N/A	2	1030000
飙车世界	1163	N/A	N/A	4360000	2270000
阿拉斯加	1120	0.32	N/A	12900000	2730000
问道的推广人ka2009	5646	N/A	N/A	23800	18
钢琴121	3940	N/A	N/A	1070000	379000
重生于康熙末年	1522	N/A	N/A		
达芬奇密码	1889	N/A	N/A	1000000	1410000
贺军翔	1942	N/A	N/A	954000	3730000

图 11-103　追词网优质关键词

你可以在这个工具中发现热门关键词，还能看到大致竞争程度，再决定自己要不要做这些关键词。

追词网工具还在不断演变中，本书第 1 版写作时（2010 年初）追词工具还是如前面所介绍的网站页面版，现在读者访问追词网站的话会发现追词工具以及改为客户端软件，功能大致相同。从他们的博客还可以看到，开发人员一直在扩展功能，修正小问题。希望这个工具以后能有更好的发展。

第 12 章

SEO 项目管理

本书的大部分章节讨论该怎样做 SEO，只要读者认真阅读，再加上一两个网站的实践，我相信一个初学者也可以很快掌握 SEO 基本技术，知道该做什么及怎么做。

但是真正做起来却很难。执行力是 SEO 行业最头疼的一个问题。在某种程度上说 SEO 没有窍门，较量的就是执行力。从我个人接触的 SEO 项目和与其他 SEO 人员交流的情况看，越大、越正规的公司，SEO 项目执行起来越困难。个人站长今天明白该做什么，明天就动手了。公司网站就麻烦得多，经常遇到这种情况，开会、写计划、讨论、批准、沟通、排期，等等，什么都没做，一个月时间就过去了。

虽然如此，SEO 人员不能怪公司大了，效率下降，这不是 SEO 部门能改变的，只能尽力推动 SEO 项目的执行。

本章就简单讨论 SEO 项目管理。在项目管理方面我是外行，相信读者中不乏公司管理和运营人员，本书讨论项目管理是班门弄斧了。所以这里只能简单讨论 SEO 工作的一些特点，以及在项目进行中可能遇到的特殊问题。

12.1 内部团队还是 SEO 服务

SEO 项目首先要确定是自己建设 SEO 团队还是使用外部 SEO 服务？

外部 SEO 服务大致上有两种。一是顾问咨询性质的专家服务，对网站进行诊断，提出优化建议或报告，但并不执行优化（如代码修改），公司内部团队需要按专家报告进行网站修改和外链建设。二是将所有 SEO 任务全部外包给服务商，包括诊断及完整执行。有的时候也可以外包一部分 SEO 任务。

对个人站长来说，通常建议不必考虑 SEO 服务，自己实践就是最好的方法。近几年我本人经常收到个人站长的 SEO 服务咨询，但通常只停留在询问阶段，无法深入。因为一般的个人网站往往连赢利模式都还没搞清楚，很难负担 SEO 服务费用。SEO 服务并不便宜，合格的诊断及优化报告一般至少也要花费几万元人民币。如果个人网站本身赢利还成问题，基本上与 SEO 服务无缘。

如果个人网站每个月赢利在几千块钱以上还需要 SEO 服务的话，要么站长应该考虑建立团队进一步扩大，要么站长本人就已经有不错的水平，不太需要第三方服务了。看

完本书，任何有建站经验的站长都可以具备基本的 SEO 水平。可以肯定，如果有第三方服务商只收几百块就能帮你做 SEO，那么他们的水平还不如你自己。

公司网站就要复杂得多，需要根据网站类型、预算、团队等现况，选择是该建立内部 SEO 团队，还是外部 SEO 服务。

1．内部团队优势

内部团队通常更了解公司各部门情况，与其他团队更能顺畅沟通，顺利执行 SEO 项目。

公司内部人员一定比外部服务商更了解本行业及公司产品，对关键词选择、文案写作、链接诱饵的设置都有优势。

公司员工更可靠、更负责，也能在最大程度上避免使用可能导致惩罚和删除的黑帽 SEO 手法。他们知道公司网站的 SEO 成败与个人利益息息相关。

2．内部团队劣势

可能需要一段时间对员工进行培训，才能达到合格的 SEO 水平。具备一定水平的 SEO 人员流动比较频繁。毕竟 SEO 是个热门行业，市场需求量不小。

潜在成本更高。一个完整的 SEO 团队，人员工资及行政成本累计起来一般比完全外包高，比顾问咨询性质的服务更高。

资深 SEO 人员难得。很多优秀的 SEO 不太愿意打工，除非公司能够提供一个非常有发展前途的平台，并且有很好的职业提升规划。很多优秀 SEO 人员要么自行创业，要么在专业 SEO 服务公司，获得的回报都不错。

3．SEO 顾问优势

立即可以使用，无须培训。

SEO 在很大程度上是依靠经验的行当。专业 SEO 顾问以 SEO 为职业，对搜索引擎及 SEO 行业新闻敏感，而且接触的网站众多，经验丰富。

专业 SEO 顾问有自己的关系网，与其他 SEO 圈内人士，甚至搜索引擎内部人员有千丝万缕的联系。在需要的时候，庞大的关系网很可能会起决定性作用。

外来的和尚好念经。在某些情况下，外部 SEO 顾问的建议更能得到公司高层的支持与各部分的配合执行，降低项目推进难度。

4．SEO 顾问劣势

缺乏对本行业及公司产品的完整深入了解。这很可能是一个致命缺陷。

在内容发布、外部链接建设、调动公司资源等方面都不可能像内部团队那样自如。

没有紧密利益联系。通常 SEO 服务商都有很多客户，无论是在责任感还是在时间分配上，外部服务都无法与内部团队相提并论。

对公司内部情况缺乏了解，需要其他部门配合时，不知道该找哪个部门哪个负责人，还是需要公司内部对应的联络人进行沟通。

可靠及安全方面的考虑，尤其是完全外包时。由外包服务商修改网站就必然要把所有数据、管理员账号、密码提供给服务商，这无疑有比较大的安全隐患。

5．怎样选择

将 SEO 服务完全外包给第三方服务商最省事，从网站修改到外部链接，从关键词研究到流量分析，服务商都有相应团队帮助完成，客户公司本身只需要设置一名联络人员进行沟通即可。但效果不一定是最好的。

我个人认为，对一个有一定规模，而且有决心做好 SEO 的公司来说，内部团队才是根本，在必要时配合外部专家的顾问咨询，或者将小部分 SEO 任务外包。只有充分了解公司内部有哪些资源可以调动、行业及竞争对手最新情况、自己的产品到底有什么不同之处、能挖掘出链接诱饵，才能使 SEO 项目真正快速高效地发挥作用。

内部 SEO 团队不需要具备最顶级的 SEO 水平，但需要紧密配合、高效运作。对 SEO 本身问题有不解之处，还可以随时使用外部顾问咨询。

12.2　寻找 SEO 服务商

网上 SEO 服务提供者多如牛毛，搜索一下“SEO”或“SEO 服务”，成千上万的公司、个人就展现在眼前。大部分网站设计公司也号称提供某种形式的 SEO 服务。寻找合适的 SEO 服务商并不是一个简单的任务。尤其在目前 SEO 这么热门，近乎泛滥的情况下，鉴别出合格的 SEO 服务商是外包 SEO 服务的关键。

1．确定外包任务

在权衡利弊之后，如果需要外包，就要明确规划外包任务及要达成的目标。在一小部分情况下，网站可能把所有 SEO 服务外包给第三方。很多时候也可以只外包一部分 SEO 任务，如只外包链接建设，或外部服务商只提供网站诊断及优化方案，或者只外包内容及软文策划和写作等。

2．寻找服务商

（1）口碑及案例

明确外包任务后，寻找 SEO 服务商最好的方法当然就是上网搜索，了解多家服务商的背景和实力，其中最主要的是口碑及案例。

就像搜索引擎排名算法本身一样，服务商自己夸耀自己的能力只是一面之词，不足为信，重要的是其他人怎么评价。参与一些 SEO 论坛、站长论坛，看看博客评论，就能大致了解一个提供商的口碑如何。

大部分 SEO 服务商与客户签订的合同中会有保密条款，客户并不希望别人知道他们做了 SEO、用的哪家服务商。所以要求备选服务商提供成功案例是一个微妙的问题。不过，一个好的 SEO 服务商应该至少有一两个案例可以公布，或者是已经获得客户同意公布的网站，或者是服务商自己的网站，有时需要潜在客户签订保密协议才能提供案例。

这里要说明的是，通过服务商本身网站在搜索“SEO”这个关键词时的排名来判断服务商实力并不是一个恰当的方法。毕竟排在第 1 页的只能有 10 个，而有实力的 SEO 服务商肯定不止 10 个。很多不错的服务商并没有把精力放在优化“SEO”这个词上，他们有其他获得客户的渠道，完全没必要在“SEO”这个词上竞争。

（2）使用什么优化方法

在初步接触阶段要求备选服务商提供优化方案不是一个好方法。要提供一个针对特定网站的实用的优化方案，需要研究市场、诊断网站，所花的时间不菲。正规的 SEO 服务商在没有确定项目时，一般不会提供完整方案。就算要求服务商提供，他们所能提供的也一定是泛泛之谈，不能用于判断服务水平。

更恰当的鉴别方式是询问一下对方大致会使用什么优化方法，从对方的答复中就可以判断是整站优化为主，还是以发垃圾链接为主，有没有使用作弊手段或者已经不起作用的优化方法。

（3）谁负责具体工作

很多 SEO 服务公司有一两个专家领队，客户之所以放心，也经常在于这一两个专家。不过在项目确定之前，应该与服务商确认项目执行时是谁或哪个团队具体负责。很有可能接洽时客户冲着核心人物的实力、口碑、名声而去，项目执行时却来了另一个项目经理或团队，甚至有可能负责日常沟通的变成了实习生。虽然不能说 SEO 服务商除了一两个人水平都不能保证，但至少应该要求对方核心人物或团队给予项目一定的关注。

（4）同业竞争条款

在接触时就应该询问备选服务商，是否正在或以前给直接竞争对手提供过 SEO 服务。如果是，最好不予考虑。

签订合同时建议加上禁止同业竞争条款，要求服务商在几年内不可以再为竞争对手网站提供 SEO 服务。这一点是十分必要的。一个服务商为同一个行业的多个公司提供服务，必然产生利益矛盾，毫无回旋余地，服务商只能为一个客户提供最好的服务。

3．小心陷阱

在寻找 SEO 服务商时需要注意避免一些陷阱。

（1）保证排名

如果有服务商保证排名能进前三页或第一页，你最好离远点。没有人能保证关键词排名，就算是搜索引擎工程师帮你优化也不能保证。谁说一定能排第一，谁就是在撒谎。

有的服务商保证的是一个毫无意义的词，比如公司名称或一个很长的没有竞争力的词，这种保证没有任何意义。

（2）声称与搜索引擎有特殊关系

如果有服务商声称自己与搜索引擎内部人员有联系，能通过私下关系搞定排名，千万不要相信。搜索引擎算法是所有搜索引擎的最高机密，能调整特定网站排名的内部人员凤毛麟角，而且这种人工调整一般是惩罚，而不是提高排名，搜索引擎不会赋予普通工程师提高特定网站排名的权力。

资深工程师或公司高层私下参与 SEO 服务的说法就更不靠谱儿了。

与搜索引擎内部人员有联系，帮助看看网站存在什么问题是可能的，靠内部关系调整排名绝不可信。

（3）不用修改网站

有的服务商会告诉你，不用修改网站就能做优化。读者看了本书前面的章节就会知道，网站结构调整至关重要，页面本身的优化也是 SEO 标准工作内容。那些号称不用修改网站就能提高关键词排名的基本上都是靠发垃圾链接。也许短时间内可以提高排名，长远来看危险大于收益。

（4）提交几千个搜索引擎

有的服务商把向搜索引擎提交作为服务的重要内容。其实向搜索引擎提交网站早就没有任何意义。

有的服务商还吹嘘可以提交几千个搜索引擎，有的夸自己使用最先进的软件提交，有的夸自己人工提交，质量有保证。不管哪种提交，都可以肯定是在忽悠。不要说几千个搜索引擎，现在有多少人能说出 10 个搜索引擎？真正有人用的搜索引擎，无论中文英文，不过四五个而已。

（5）优化谁的网站

有的服务商并不优化客户网站，而是建立和优化新网站，获得流量后通过链接或转向把访问者引到客户网站。

本质上说，这是在卖流量，而不是在提供 SEO 服务。尤其是如果优化的网站是由服务商本身所拥有或控制的，是非常危险的，一旦有任何争议，服务商可以立即切断流量，甚至把流量卖给你的竞争对手。

（6）发垃圾邮件

如果你是通过垃圾邮件知道某家 SEO 服务商，最好躲得远远的。与此类似，打骚扰电话推销 SEO 服务的一般也不会是好的提供商。这样的 SEO 服务商自己还在发垃圾邮件呢，他们怎么可能给你带来排名和流量？

（7）与广告混为一谈

这是保证排名的另一种形式。很多人不一定分得清楚搜索结果页面上自然排名与广告的区别。有的服务商利用这一点向客户保证排名出现在第一页，其实是出现在第一页的广告中。当然不是说搜索广告不能做，而是正规的 SEO 服务商不能把自然排名与广告混为一谈。这两者之间没有任何关系，做广告就按广告的方法做，付广告的点击费用。做 SEO，就要做更长远打算，停止付广告费，排名还会保持在那里。

12.3 SEO 团队建设

一个正规完整的 SEO 团队应该包括以下人员。

1. SEO 经理

负责公司 SEO 统筹及管理。具体工作包括：

- SEO 目标的制定以及 SEO 整体策略规划，包括内容及链接策略等。
- 统筹和沟通，既包括与其他部门的沟通协作，也包括 SEO 部门内部任务及计划的制订和执行。
- 竞争对手和关键词分析。

- 网站架构设计。
- 标准制定，对公司网站的新建内容、HTML 代码、页面优化都应该有内部规范。
- 培训，SEO 团队内部的培训提高，以及帮助其他部门了解 SEO 基本常识的培训。

SEO 经理需要具备比较高的水平，掌握搜索引擎和 SEO 原理及方法，具备比较多的市场营销知识和经验，不能仅仅局限于 SEO 技术，也要了解网站建设、HTML 代码，基本的编程知识，管理和沟通技巧，对 SEO 行业动向足够敏感，并且与 SEO 行业人士有比较多的联系沟通。

SEO 经理应该有足够大的权利。在很多公司，SEO 人员地位过低，提出优化建议却不能调动其他部门，尤其是技术部门，使计划很难实施。SEO 经理通常至少应该是中上层管理人员，获得高层充分授权。

2．页面优化人员

负责人工调整页面优化因素，包括人工调整重要页面标题、正文内容、关键词分布及格式，增加和调整 tags，调整内部链接等。

3．内容编辑

原创或编辑网站内容，发布内容时做基本的 SEO 优化，比如标题的撰写、关键词研究及专题组织。用户提供内容的网站也可能需要有内容编辑进行审核及修改。链接诱饵的设置，尤其是资源型链接诱饵，也是内容建设人员的重要工作。

4．链接建设人员

友情链接交换，依托于有价值内容的链接请求，社会化媒体网站上的账号创建、管理及内容发布，文章发布，论坛及博客参与和评论，设计和制作链接诱饵等。

5．技术及设计人员

根据 SEO 经理设定的网站架构，对网站内部结构、分类设置、URL 规范化及转向进行调整。自动生成的页面进行代码及关键词优化。技术在大部分公司是单独的部门，SEO 人员需要与其紧密合作。

6．流量分析

记录、统计搜索流量，找出存在的问题，发现新的流量来源，提供 SEO 策略修正方案。

SEO 团队可大可小，根据公司规模及对 SEO 的投入，既可以是一个人，也可能多达几十人。目前优秀的 SEO 人才流动性较高，虽然有点石论坛这样的专业平台及各种招聘服务，但寻找到合适的 SEO 并不是一件容易的事。从公司内部培训市场营销或技术人员充实 SEO 团队是更可行的方法。

12.4 流程及计划

SEO 工作细小繁杂，千头万绪，又经常涉及多个部门，如果不事先计划好，往往容易无从下手，或者这里做一点那里做一点，失去整体方向。SEO 团队负责人应该计划好未来 6 个月或 12 个月内，应该完成哪些工作，怎样完成，由谁完成。

1．记录所有修改

除了记录效果监测一章提到的收录、搜索流量、指标关键词排名、转化等数据外，还有一个至关重要却经常被忽略的数据需要完整记录，那就是网站上的所有修改。有的修改与 SEO 直接相关，可能就是为了 SEO 而做的，有的是因为其他种种原因，由不同部门所做，却可能影响 SEO。

一般来说，网站优化是个长期过程，即使不考虑新栏目、内容、产品，现有网站结构和页面元素都可能不断修正。如果没有完整记录网站上的改动，过一段时间排名有变化，SEO 人员将无法知道到底为什么变化。

虽然由于搜索引擎排名算法的复杂性，网站修改与排名和流量之间不能因为时间关系而轻易确定因果联系，但至少能让 SEO 人员有个分析的起点，找到大致可能的原因。没有记录网站变化历史，就基本两眼一摸黑了。

应该记录的网站修改除了列出详细改动内容及原因，还要列出准确时间。

2．设定工作目标

这里所说的目标不是指 SEO 总体目标，如排名和流量，而是指具体的、SEO 人员可以实施的工作。如：

- 一年内增加外部链接 1 万个。
- 一年内完成页面标题、H 标签、图片 ALT 文字的所有优化。
- 2 个月内完成导航系统优化。
- 6 个月内完成首页、分类页面及主要产品页面标题人工关键词研究、优化撰写。
- 6 个月内收录达到 5 万等。

SEO 工作目标的设定，要分析自身资源及竞争对手情况，不能设下不切实际的目标，

如一年内增加外部链接 100 万，但人员、预算限制使这种目标不可能达到。如果主要竞争对手外部链接都在几万水平，你要建设 100 万个外部链接也完全没有必要。

3．任务分解

设定好 SEO 工作目标后，还可以进一步细化和分解 SEO 任务。比如要达到一年内增加 1 万个外部链接，可以分解为：

- 友情交换 300 个外部链接。
- 制作 WP 模板 5 个。
- 写作资源型链接诱饵文章 20 篇。
- 目录提交成功 50 个。
- 创建免费博客 20 个，各发表帖子 30 篇以上。
- 网站主博客发表帖子 50 篇以上。
- 联系行业内权威博主 20 人，密切沟通。
- 发表新闻稿 5 篇。
- 购买链接 20 个。

再比如要达到收录 10 万的目标，任务可以分解为：

- 内容编辑人员确保实际产品或内容页面达到 20 万。
- 一年内至少创建 10 个专题栏目。
- 项目经理两个星期内完成网站结构及导航最终设计。
- 技术部门两个月内按 SEO 提供的方案完成网站结构调整及 URL 静态化。
- 网页优化人员人工调整部分重要内部链接。

任务分解后才能更清楚是否有完成的可能，需要怎样分配人员和资源，多长时间能完成。

4．人员责任

SEO 任务分解后，每一项具体任务都要分配到特定人员负责。根据 SEO 人员自身特长分配明确具体的任务，不仅有助于提高效率，也方便监控项目进度、明确责任。

5．时间资源分配

列出详细 SEO 任务及负责的具体人员后，还要规划这些任务应该在多长时间内完成，以及其他资源如预算的分配。以外部链接建设为例，可以参考如表 12-1 所示的 SEO 计划分配。

表 12-1　SEO 计划分配

	2012 年 1～3 月					2012 年 4～6 月						2012 年 7～9 月					2012 年 10～12 月			
友情链接	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成
	75	T1			70	75	T1			90	75	T1			75	75	T1			80
WP 模板	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成
	2	T2		技术部	1	1	T2		技术部	2	2	T2		技术部	2	1	T2		技术部	1
目录提交	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成
	20	T1	1000		20	10	T1	500		8	10	T1	500		8	10	T1	500		10
新闻稿	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成
	3	H	3000	公关部	1	1	H	1000	公关部	2	2	H	2000	公关部	2	1	H	1000	公关部	2
诱饵文章	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成	指标	负责人	预算	支援	完成
	5	T3		编辑部	10	5	T3		编辑部	5	5	T3		编辑部	8	5	T3		编辑部	5

6．流程定义和规范

每一项 SEO 任务都应该有比较明确的规范，不仅有助于提高整体 SEO 工作质量，也能使新加入的人员快速熟悉应该怎样完成自己的任务。

这些规范包括：

- 人工撰写标题时，怎样进行关键词研究及标题写作规范。
- 怎样判断潜在链接来源的质量和价值。
- 交换友情链接时使用的邮件模板，以及应该注意的事项。
- 创建专题内容时怎样监测热门关键词。
- 怎样收集和编辑内容。
- 资源型链接诱饵概念创意过程及写作方式。
- WP 模板制作规格以及推广场所、渠道。
- 与其他博主沟通留言时的注意事项。

这些规范都可以只是很简单的几条，但明确写出来才能让 SEO 人员时刻注意，不能降低 SEO 工作标准。

12.5 绩效考核

SEO 部门和工作人员绩效考核方法是个难点，不少公司都在探索完善中。

网站多少个关键词在一段时间内排名达到前一页或前两页，或者搜索流量在几个月内达到多少日 IP，看似是好的绩效考核标准，但其实不容易准确反映出 SEO 人员的工作量。排名和搜索流量并不受 SEO 人员控制，就算 SEO 部门做了大量的工作，排名和流量也可能没有显著提升。仅仅因为排名流量不理想就否定 SEO 部门的工作显然不公平，也不能激发 SEO 人员的工作热情，不能让 SEO 人员明确今后的工作方向。

所以排名和搜索流量可以作为整个 SEO 部门绩效考核的一部分，但针对每个 SEO 人员的绩效考核，建议加入真实反映工作量的内容。这种绩效考核目标需要是具体的和可测量的。

举几个例子。编辑部门绩效考核指标可以是：

- 每天原创文章一篇。
- 每星期编辑一百篇转载内容。
- 每星期撰写一篇资源型链接诱饵内容。
- 每天编辑、调整 10 个客户发布的 B2B 产品信息。
- 每个月创建 3 个热门搜索词的专题栏目。

外部链接建设可以设置这样的指标：

- 每星期联系 100 个潜在友情链接网站。
- 每星期确保获得 20 个友情链接。
- 每天利用高级搜索指令，找到 100 个可以联系的外部链接来源网站。
- 每星期找到并发出 10 个单向链接请求邮件。
- 每天在相关论坛发帖或回帖 10 篇。
- 每个月设计 WP 模板 2 个。

从上面举的例子可以看到，衡量工作量的指标必须是某种具体行动，而且有明确数目可以衡量。只要 SEO 策略、方向正确，这种具体工作量长期累积下来，一定会导致排名与搜索流量的提高。在对 SEO 人员绩效考核过程中，具体可测量的工作量应该占主要部分。

12.6 获得高层支持

在大中型公司，没有公司高层的全力支持，SEO 经常寸步难行。有一定规模的公司，SEO 人员很少是最终执行网站优化的部门。SEO 报告和计划通常需要交由技术部门、内容编辑和产品部门等执行。SEO 部门在公司里的位置最多是与其他部门平级，加上其他部门人员可能对 SEO 没有什么概念，SEO 人员要想顺利推动其他部门执行优化，是一件很有挑战性的工作。其他部门各有自己份内的工作，SEO 人员提交的优化建议需要排期，连一个很小的改动也要等上几个月并不罕见。来自高层的指示和压力就将非常重要，往往具有决定性意义。

1．SEO 价值

公司高层不会对 SEO 技术细节感兴趣，但是会对 SEO 带来的价值与好处很感兴趣。SEO 人员为说服高层，可以在前期关键词研究的基础上，以明确详细的数据及图表说明 SEO 对公司营收的价值。

比如使用下面这类数据。

公司业务最主要的 3 个关键词，在百度和谷歌月搜索总次数为 20000 次。

根据搜索结果前十名点击量分布，排在第一名的结果将获得 42%的点击。如果 3 个关键词都排在第一，将获得 8400 个搜索流量。如果排在第五名，点击率为 4.9%，也将获得 980 个流量。

经过关键词研究，网站分类页面设定二级关键词 50 个，总搜索次数为 50000。如果都排在第一位，将获得 21000 个点击。排在第五位将获得 2450 个点击。假设 50 个关键词排名成功率 50%，最好获得 10500 个点击，最坏也将获得 1225 个点击。

如果 SEO 较为成功，一级和二级关键词将获得 18900 个点击。

通常大中型网站长尾流量占 60%～70%，所以一个成功执行 SEO 的网站，首页、分类页面加上长尾流量，将获得 63000 个点击流量。

基于网站现有转化率为 1%，63000 个流量将带来 630 个订单。网站平均每单销售额为 100 元，搜索流量将带来 63000 元订单。网站历史数据显示平均利润率为 30%，上述搜索流量将带来 18900 元毛利。

所以在成功实施 SEO 后，可以为网站每月带来 63000 流量，以及 63000 元销售，18900 元毛利。

公司高层可以对技术细节不感兴趣，但是对具体的流量及收入增长潜力一目了然。

当然，再成功的 SEO 执行也不可能把所有关键词都做到第一。SEO 部门应该根据自身人员、资源投入、竞争对手情况，估计可以实现的目标，列出最好情况下的 SEO 价值，以及正常情况下的 SEO 价值。

2．竞争对手情况

有时候指出竞争对手排名和搜索流量的成功，自己因为没有 SEO 而失去的市场机会，能带给高层更大紧迫感。SEO 部门可以在报告中列出主要竞争对手目前关键词排名情况，根据 Google 趋势等流量工具得到的流量数字，竞争对手流量占据总搜索次数的市场份额，竞争对手在 SEO 方面做了哪些工作等，高层人员对同行业整体市场份额分布将有更清晰的了解，对自身网站可能的增长潜力也会有更清晰的预期。

3．预算

天下没有白吃的午餐，实现 SEO 价值不可能没有代价。SEO 部门需要根据前面 SEO 流程中提到的任务分解、时间资源分配，计算出未来一段时间需要投入的人员、资金，以及需要其他部门配合的地方，计算出 SEO 项目的预算。

4．设定期望值

SEO 带来的收入和赢利潜力，通常会是个很让人期待的数字。不过还要给高层设定一个现实的期望值。不要让老板觉得经过 SEO，流量和销售唾手可得，板上钉钉，以免今后期望越大，失望越大。

提交给高层的报告中应该明确指出存在的风险和代价，除了预算外，还包括执行 SEO 需要比较长的时间，通常至少 6～12 个月，不要把 SEO 与搜索竞价相提并论。SEO 成功率不是自己能够决定的，就算 SEO 策略、方法完全正确，还必须考虑到搜索引擎算法的改变，其他部门的配合等情况。

5．执行报告

为持续获得高层支持，定期提交 SEO 报告是必要的。首先需要确定流量分析以及跟

踪主要关键词排名所使用的软件或工具，一旦确定后不要轻易更改。每种工具都有自己的局限性，不同工具在同一时间记录的数据很可能不同。但只要持续使用相同工具，各项数据随时间变化的趋势才是最重要的。

然后设定报告中应包含的各项 SEO 目标，包括收录数、各级关键词排名、总搜索流量、转化率、搜索流量带来的销售收入等。

为了准确显示 SEO 效果，必须设定比较基准，包括两部分：一是 SEO 执行前网站的收录、排名、流量等数字，二是主要竞争对手的收录、排名、流量。

每个月甚至每星期，固定时间记录各项指标变化情况。最好能以图表形式显示出随时间推移，收录、排名与流量相对于最初基准的增长或下降。

完善的执行报告既能向高层汇报 SEO 的效果，也能让 SEO 部门本身明确工作进程和成效。看到各项指标数字的增长，日常看似无聊的内容编辑、链接建设，都将变得有意义。

12.7 沟通、培训及规范

由于各部门任务分配交叉，很多优化工作并不是 SEO 人员最终直接完成的。公司越大，越依靠其他部门执行 SEO 方案。SEO 部门经常处于策划、建议、监督的位置，代码的修改需要技术部门完成，网站栏目规划、创建需要产品部门完成，撰写和编辑内容需要编辑部门完成。

1．沟通

SEO 部门有责任使公司全体部门意识到 SEO 的价值，SEO 的成功与否关系到整个公司和网站的成败。没有流量，程序、内容都将失去意义（当然，SEO 只是网络营销、吸引流量的其中一个方法）。

与其他部门的积极沟通十分重要，SEO 项目经理的重要工作内容之一就是沟通。沟通技巧与 SEO 技术本身关系不大，不再详细讨论。唯一的建议是，SEO 部门可以设置一个 SEO 进度内部公告板，供公司全体查看。公告板上贴出 SEO 工作目标以及分解后的详细任务，以及时间、责任分配表，明确哪些工作应该由哪个部门在哪段时间完成，标注工作完成进度，没有完成的工作，高层及全体人员都可以从公告板中看到责任在哪个部门，督促公司全体人员按计划执行设定好的 SEO 任务。

2．内部培训

内部培训也是 SEO 部门重要工作内容之一。其他部门人员对 SEO 细节毕竟不能完全理解，对 SEO 任务中的具体更改很可能知其然而不知其所以然。我在顾问咨询服务过

程中就经常遇到这种情况，技术部门人员没能充分理解优化细节的深层意义，自由发挥使用他们自己认为是同样效果的其他方法，其实与 SEO 要达到的目的南辕北辙。

要确保其他部门能准确执行 SEO 计划，SEO 部门最好能够定期组织内部 SEO 培训，讲解 SEO 基本原理以及最新趋势，尤其是技术、产品和内容编辑部门。对大中型公司来说，SEO 培训几乎是必需的。

3．建立规范

建立网站 SEO 规范也是 SEO 部门的重要工作内容。互联网公司人员流动频繁，新来的员工很可能对 SEO 一无所知。给各部门发布需要遵守的 SEO 规范，是保证网站不出现重大技术失误的重要方法。SEO 项目经理应该为各部门制定简单而明确的技术规范，规范文件不需要很长，几页纸就能防止网站上出现重大搜索引擎不友好内容。

比如给技术部门的规范可以包括：

- URL 命名系统规范。
- 系统自动生成页面 Title、H1、图片 ALT 文字规范和格式。
- 网站地图生成规范。
- 面包屑导航规范。
- CSS、JS 脚本使用规范。
- 网站内部链接及栏目设置规范。

给内容编辑部门的 SEO 规范可以包括：

- 关键词研究流程。
- 文章标题写作规范和举例。
- 文章关键词分布要求。
- 正文内部链接生成方法及规范。
- 监测热门关键词及建立专题流程。
- 用户贡献内容审核编辑规范。

大中型网站的 SEO 不是 SEO 部门或人员本身能完全控制的，必须依靠其他部门的配合才能达到最好效果。

12.8 应急计划

虽然谁都不希望 SEO 项目出现问题，但搜索引擎排名和流量不是站长自己能控制的，出问题是非常常见和正常的，应该提早做准备。商业网站更应该有适当的应急计划。

1．监控

应急计划的第一部分是随时监控。SEO 出现问题的常见现象主要有两方面，一是总

搜索流量下降，二是网站主要关键词排名大幅下降。

搜索流量周期性波动很正常，比如周末流量通常会比工作日低很多，有的网站可能会低 30%～40%。但同样在工作日，如果总搜索量下降 20%以上，说明很可能出现了问题。

一般来说，单个关键词排名波动也很正常，尤其对大中型网站来说，一部分关键词排名上升，另一部分排名下降是正常现象，只要总搜索流量稳定通常不是问题。

网站核心关键词排名在 20 名以内波动，一般也是正常现象。但如果突然下降三四页以上，也说明很可能 SEO 方面有问题，网站被降权。

效果监测章节所提到的日常监测内容将对判断和启用应急措施起到至关重要的作用。网站如果同时出现下面三种情况，应该启动应急措施：

- 总搜索流量下降 20%以上。
- 首页主关键词下降三四十位。
- 所有长尾关键词排名整体明显下降。

2．替代流量

个人网站和信息类网站搜索流量下降一段时间，一般还可以承受，最多是收入减少。但对某些网站来说，搜索流量明显下降很可能给公司运营带来致命性的打击，尤其是完全依靠网站进行销售的电子商务网站。流量大幅降低，客服、物流等运营成本不会下降，为维持公司运作，必须立即启动替代流量来源，哪怕不赚钱，至少维持现金流。

最方便的替代流量来源是搜索广告，只要开通账号，设置相应关键词，立即会带来流量，而且流量质量与免费搜索流量不相上下。搜索广告包括百度推广、Google AdWords 等。搜索广告的设置和管理，请参考《Google 广告优化与工具》、《Google Adwords 营销——网商成功之道》等书。

另外一个马上会有效果的替代流量是网站自身积累的邮件列表。无论是免费电子杂志的订户，还是数据库中的付费用户，这时都可能成为最好的暂时替代流量来源。只要向数据库中的电子邮件地址发几封推广邮件，就可以带来不少重复流量，把以前的用户拉回网站。电子邮件营销是最有效的网络营销方法之一，具体操作方法可以参考《网络营销实战密码》一书的电子邮件营销章节。

网络广告、事件营销、论坛营销等也都可以作为替代流量来源选项。

3．网站诊断

确定 SEO 方面出现问题，启动替代流量的同时，要立即对网站进行问题诊断。除了黑帽与惩罚章节中提到的诊断过程和方法外，还要注意以下几点。

在监控自己网站的同时，也要注意主要竞争对手网站排名情况，是否他们的整体排名都有下降？如果以前排名不错的主要竞争对手排名都下降，很可能说明问题在于搜索引擎调整算法。那么掉下去的网站有没有共同特征？新上来的网页又有什么共同特征？

关注搜索及 SEO 行业最新趋势。与行业中其他从业者的沟通联系此时就显得很重要。当搜索引擎算法有重大变化时，行业人士通常会在博客、论坛中讨论，提出可能的原因以及补救方法。有的时候搜索引擎算法改变效果不如预期，过几天还可能再变回去。多听多看，先不要轻举妄动。

关注关键词搜索趋势。使用 Google 趋势、百度指数等关键词工具查看搜索流量下降是否因为用户关注度及搜索次数下降。如果整个产品线或行业关注度都下降，就不是网站自身努力所能补救的了。

仔细查看记录的网站修改日志。前面提到的网站修改日志对诊断 SEO 问题至关重要。如果只是自己的网站出现问题，很可能是在这之前某个时间的网站修改方向错误，分析哪些修改最有可能导致问题，是否有过度优化之嫌，尝试逐步修改回原始状态。

有的时候网站搜索流量整体下降并不能找到具体原因。这时只能依靠增加原创内容，继续吸引高质量外部链接，积累域名权重，等待搜索引擎重新评估网站。

第 13 章

搜索引擎排名因素调查

前面的章节讨论了很多 SEO 技术。本章分别列出 Google 和百度的一百多个排名因素，并且给予重要程度评分。读者可以把这两个排名因素列表当做速查手册使用。列出的排名因素在本书中都有详细介绍，读者可以参考对应章节的内容。

这两个排名因素列表及评分都是通过对 SEO 领域的专家问卷调查所得的，代表了集体智慧，不是我本人的观点，因而更加准确。

13.1 Google 排名因素调查 2009

本节是 SEOmoz 2009 年度的搜索引擎排名因素调查。SEOmoz 正式授权在本书中发表中文翻译版。英文原文地址：

http://www.seomoz.org/article/search-ranking-factors

虽然笼统称为“搜索引擎排名因素”，但调查实际上是以 Google 为基础的。Google 中英文网站排名使用的是同样的数据库和核心算法，所以对中文网站在 Google 的排名也有很大参考意义。

> 每两年 SEOmoz 都对全世界 SEO 领域的专家进行问卷调查，询问搜索引擎排名算法方面的意见。每个参与者都收到 100 多个排名因素及 SEO 领域的一些热门问题。这份文件代表了参与专家的集体智慧，我认为是最重要的 SEO 资源之一，为大家提供了搜索引擎优化在实践中哪些因素有用，哪些没有用。
>
> Rand Fishkin,
> SEOmoz 创始人、CEO

说明：

下面的排名因素由 72 位 SEO 专家评分，他们的反馈意见整合为下面所显示的百分比。每一个因素我们都计算出专家们所认为的获得搜索引擎排名的重要程度，以及根据回答的标准差得到的争议度。也就是说，高重要度加上低争议度，代表着大部分专家都同意是搜索引擎排名的关键因素。

重要程度如表 13-1 所示。

表 13-1　主要程序

百 分 比	代表的重要程度
65%～100%	非常重要
55%～64%	很重要
45%～54%	中等重要
35%～44%	重要度比较低
25%～34%	不太重要
0～24%	很不重要

争议程度如表 13-2 所示。

表 13-2　争议程度

百 分 比	代表的争议程度
0～3.9%	高度共识
4%～7.9%	中度共识
8%～11.9%	轻度共识
12%～15.9%	中度争议
16%～20%	高度争议

总体排名算法如图 13-1 所示。

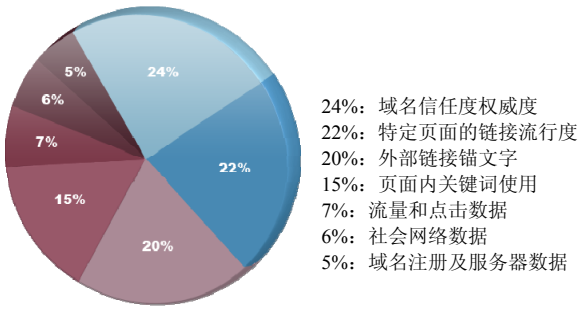


图 13-1　2009 年 Google 排名因素比重

5 大正面排名因素如表 13-3 所示。

表 13-3　5 大正面排名因素

排　　名	因 素 说 明	重 要 程 度
1	外部链接锚文字包含关键词	73%，非常重要
2	外部链接流行度（外部链接的数量质量）	71%，非常重要
3	链接来源多样性（链接来自于很多不同域名）	67%，非常重要
4	标题标签中任何地方使用关键词	66%，非常重要
5	基于与被信任域名的链接距离得到的域名信任度（例如 TrustRank 算法）	66%，非常重要

5 大负面排名因素如表 13-4 所示。

表 13-4　5 大负面排名因素

排　　名	因 素 说 明	重 要 程 度
1	恶意操纵为目的的隐藏页面	68%，非常重要
2	从有名的链接买卖提供商那里得到的链接	56%，很重要
3	页面上有连向垃圾网站或页面的链接	51%，中等重要
4	检测用户代理类型（浏览器还是蜘蛛），然后隐藏页面	51%，中等重要
5	服务器经常死机，网站不能打开	51%，中等重要

5 大最具争议因素如表 13-5 所示。

表 13-5　5 大最具争议因素

排　　名	因 素 说 明	争 议 程 度
1	检测是否支持 Cookie，然后隐藏页面	16.3%，非常争议
2	检测 JS、富媒体支持，然后隐藏页面	15.4%，中度争议
3	与背景相同或相似颜色的隐藏文字	15.3%，中度争议
4	检测访问 IP 地址，然后隐藏页面	15.3%，中度争议
5	检测用户代理类型（浏览器还是蜘蛛），然后隐藏页面	15.2%，中度争议

有共识或有争议的百分比是根据被访者回答的标准差计算的。

被访人根据 Google 排名算法回答问题。虽然我们发现这些回答和数据也适用于其他主流搜索引擎（如 Bing，雅虎，Ask）等，当然某些不同是肯定存在的。

13.1.1　与关键词有关的页面排名因素

与关键词有关的页面排名因素如表 13-6 所示。

表 13-6　与关键词有关的页面排名因素

排　　名	因 素 说 明	重 要 度	争 议 度
1	标题标签中任何地方使用关键词	66.00%	8.00%
2	标题标签中第一个字或词使用关键词	63.00%	11.30%
3	域名中包含关键词	60.00%	11.20%
4	H1 标签中任何地方使用关键词	49.00%	10.20%
5	页面上导出内部链接锚文字中使用关键词	47.00%	13.00%
6	页面上导出外部链接锚文字中使用关键词	46.00%	13.60%
7	H1 标签中第一个字或词使用关键词	45.00%	11.70%

续表

排　名	因素说明	重要度	争议度
8	页面最前面 50～100 个可见文字中使用关键词	45.00%	9.90%
9	子域名中包含关键词	42.00%	9.00%
10	文件名中包含关键词	38.00%	9.10%
11	目录名中包含关键词	37.00%	8.60%
12	其他正文标题（H2 到 H6）中使用关键词	35.00%	8.00%
13	图片 ALT 文字中使用关键词	33.00%	8.70%
14	页面文字中关键词使用和重复次数	33.00%	10.30%
15	页面上使用的图片文件名中包含关键词	33.00%	8.60%
16	黑体（<b>或）中使用关键词	26.00%	7.60%
17	关键词密度（关键词使用次数除以页面上总单词数）	25.00%	9.80%
18	页面上列表（ ）中出现关键词	23.00%	9.50%
19	页面 URL 参数中出现关键词（如 seomoz.org/page.html?keyword）	22.00%	7.60%
20	斜体（<i>或）中使用关键词	21.00%	8.40%
21	描述标签中使用关键词	19.00%	9.90%
22	页面文件扩展名中出现关键词（如 seomoz.org/page.keyword）	12.00%	8.30%
23	HTML 注释标签中使用关键词	6.00%	5.70%
24	关键词标签中使用关键词	5.00%	5.50%

与关键词有关的页面排名因素的评论如下。

- Andy Beal：总体上来说，外部链接锚文字中使用关键词是最重要 SEO 因素之一。我看到一些网站页面上根本没有出现关键词，但是在一些竞争很强的关键词排名中表现很好，仅仅是因为外部链接锚文字的因素。
- Andy Beard：可以忽略在关键词标签中使用关键词，除非博客系统显示标签时要用到这些关键词。Google 完全忽略关键词标签。
- Christine Churchill：花时间写一个好的标题标签，这是所有页面优化中回报率最高的。好好写标题标签。
- Duncan Morris：值得指出的是，描述标签中出现关键词并不影响排名，但是会直接影响搜索引擎排名结果的点击率。
- Peter Wailes：域名中包含关键词之所以有一定的效果，是因为很多人链接向你时，只使用域名本身作为锚文字。域名本身的价值很有限。

13.1.2　与关键词无关的页面排名因素

- 与关键词无关的页面排名因素如表 13-7 所示。

表 13-7　与关键词无关的页面排名因素

排　　名	因 素 说 明	重 要 度	争 议 度
1	独特的实质性的原创内容	65.00%	9.20%
2	页面创建的新鲜度	50.00%	10.50%
3	页面上有链接连到同一个域名下的其他 URL	41.00%	12.60%
4	页面内容历史变化（页面内容更新多频繁）	39.00%	10.90%
5	页面上有指向其他网站的链接	37.00%	13.30%
6	包含参数的是动态 URL 还是静态 URL	33.00%	11.80%
7	HTML 中格式代码与可见文字之比	25.00%	11.00%
8	使用描述标签	22.00%	11.00%
9	W3C 标准验证	16.00%	9.30%
10	使用 Flash（或其他需要插件的内容）	13.00%	10.10%
11	页面上有广告	11.00%	8.60%
12	页面上有 Google Adsense	8.00%	7.30%

与关键词无关的页面排名因素的评论如下。

- Russell Jones：如果 Google 只给那些有历史的页面好排名，它们的排名结果将很陈旧。在链接很难得到的时候，新鲜度是一个很有价值的资产。
- Tom Critchlow：新鲜度和内容更新这种因素很难确认。如果你要排名的是很新的搜索词，当 QDF 起作用时，那么新鲜的页面是有价值的资产。但其他情况下，一个有长远历史的页面更有利。所以有的时候你需要这个，有的时候你就需要那个。
- Peter Meyers：新鲜度比任何时候都更重要。我经常惊讶的看到，一个博客帖子一天之内就能出现在第一页，然后过一段时间稳定在低几个位置或几个页面的排名上。
- Carlos Del Rio：HTML 验证并不是必要的，但是验证一下页面代码是发现错误代码、避免蜘蛛陷阱的简单方法。如果你完全没有连到其他地方的链接，只是一个信号，代表你的网站不是整个网络的一部分。建立话题性的关联是维持好排名的重要因素。
- Ian Lurie：格式代码与文字比例及 HTML 验证并不直接影响排名，但是这些因素可以帮助你建立正确的语义标志及快速下载有丰富内容的页面，而这些就有很大的影响了。说明标签和静态或非静态 URL 不影响排名，但是它们影响点击率。所以我倒不是建议忽略描述标签或使用看起来很乱的 URL。但是如果你改进它们，你会看到同样排名下的更多点击，倒不是更好的排名。

13.1.3　特定页面链接流行度排名因素

特定页面链接流行度排名因素如表 13-8 所示。

表 13-8　特定页面链接流行度排名因素

排　　名	因 素 说 明	重 要 度	争 议 度
1	外部链接锚文字出现关键词	73.00%	6.40%
2	外部链接流行度（外部链接数量质量）	71.00%	9.20%
3	链接来源多元化（链接来自很多不同域名）	67.00%	8.50%
4	特定页面的 TrustRank（这个页面是否从被信任的网站获得了链接）	65.00%	8.70%
5	基于迭代算法得到的整体链接流行度（例如 Google PR）	63.00%	8.80%
6	外部链接来源聚焦于特定主题（页面的外部链接是否来自与主题相关的页面和网站）	58.00%	10.60%
7	内部导入链接锚文字使用关键词	55.00%	9.90%
8	页面在网站结构中的位置（页面在网站结构关系图中处于什么样的位置）	51.00%	10.70%
9	内部链接流行度（来自同一个域名下其他页面的反向链接）	51.00%	9.10%
10	NoFollow 链接的数量质量	25.00%	10.80%
11	连到页面的 Follow 和 NoFollow 链接的比例	17.00%	11.40%

特定页面链接流行度排名因素的评论如下。

Jon Myers：对我来说，现在的 SEO 排名赢就赢在外部因素上。还是那个古老的 80/20 原则，你必须花时间获得正确的链接才能赢。确保使用关键词做锚文字，并且把链接导向相关页面。重点是质量和数量的平衡混合，不能所有链接来自于一种来源，要确保来源是混合在一起的。

使用 PR 值作为判断好链接的因素，但同时也要确保相关性。质量高的链接更要使用关键词做锚文字，我发现 85%以上的外部链接锚文字都是品牌名称，所以你可能失去了很多好机会。另外别忘了内部链接也很重要。

Russell Jones：链接为王，链接万岁。

Hamlet Batista：优化一般但有很多反向链接的页面，很轻松地就能超过优化很好但链接很差的对手。

Todd Malicoat：链接对 SEO 来说，就像雪对爱斯基摩人一样。页面之外的因素是搜索相关性方面最大的改变，也使 Google 成为今天这样的巨人。在这方面多花心思，理解不同链接的区别及其和搜索结果的关系，你就理解了 SEO 的核心。明白怎样评估一个网站的链接资产，你就具备了一套在搜索结果中评估竞争对手的强有力的技能。

Jane Copland：我当然不认为 Follow 和 NoFollow 链接比例对传统 Google 排名结果有很大影响。想想诸如 Twitter 这样的高质量社会网络链接，都是有 NF 标签的。把大量 NF 链接当做不利于信任度的衡量标准是件荒谬的事。

13.1.4 全站链接排名因素

全站链接排名因素如表 13-9 所示。

表 13-9 全站链接排名因素

排 名	因 素 说 明	重 要 度	争 议 度
1	基于与被信任域名链接距离的域名信任度（如 TrustRank 等）	66.00%	9.50%
2	基于迭代算法得到的域名整体链接流行度（域名 PR 值等）	64.00%	11.00%
3	域名链接多元化（指向这个域名所有链接的域名数量及种类）	64.00%	9.50%
4	来自特定主题的枢纽/权威网站的链接（Hilltop 算法）	64.00%	10.90%
5	指向域名的链接短时间内快速增长或减少（一段时间内获得链接的数量、质量及时间轴上的发布）	52.00%	9.50%
6	来自不能随便注册的域名的链接（如.edu，.gov，.mil 等）	47.00%	13.80%
7	指向域名的 Follow 和 NoFollow 链接比例	21.00%	11.00%

全站链接排名因素的评论如下。

Carlos Del Rio：很可能存在一个 NoFollow 链接与 Follow 链接之比的引爆点。没到这个点之前，它不是排名因素。但 NF 链接过多，达到这个点之后，就变成一个负面影响因素。

Will Critchlow：从我的经验看，超过本身价值的链接突然增长，只对 QDF 类型的查询有正面影响。

Aidan Beanland：Google 过去曾经指出过，.edu .mil .ac 等域名本身并不比其他域名更有价值。但是其他因素的交互作用，使这些域名好像更有价值。

Ann Smarty：域名权重现在还是一个非常重要的因素。我们经常看到一些权重很高的域名上的页面，在搜索竞争非常强的词时排名很好。

Lisa D Myers：我认为与被信任域名的远近确实有影响，搜索引擎蜘蛛正在变得更智能化，读取网页时会比较域名之间的关系。LSI 对搜索引擎来说只是一个开始，我相信现在的算法已经更复杂，不仅能够读取一个页面上内容之间的语义，也能读取网站之间的语义关系。这种复杂的处理过程很难想象，就好像你小的时候尝试想象宇宙的终点那样复杂。不过最后还是回到内容，如果你能产生高度有价值并相关的内容，好的链接自然会来。我知道这是个老生常谈，但不幸的是，这是事实。如果链接是网上的钞票，内容就是银行。

13.1.5 全站非链接排名因素

全站非链接排名因素如表 13-10 所示。

表 13-10　全站非链接排名因素

排　　名	因 素 说 明	重 要 度	争 议 度
1	网站架构（是否使用了清晰有用的结构）	52.00%	13.00%
2	链接到其他被信任、有权重的网站或网页	37.00%	10.80%
3	域名注册年数	37.00%	14.30%
4	域名注册历史（域名已经注册在同一个所有人名下多久？续费次数等）	36.00%	12.30%
5	主机在线率	32.00%	11.40%
6	主机信息（同一台主机和 IP 地址上还有哪些域名？）	31.00%	10.40%
7	域名所有人的改变（根据注册信息，域名是否曾经转手？）	31.00%	11.30%
8	Google 资讯中有这个域名的 Feeds	31.00%	14.90%
9	使用 XML 网站地图	29.00%	12.30%
10	域名所有人（谁注册了域名？以及所有人的历史）	25.00%	12.10%
11	注册 Google 本地服务	24.00%	12.70%
12	域名被提到情况（域名和网址以文字形式被提到，但不是链接）	24.00%	9.80%
13	Google 博客搜索中有这个域名的 Feeds	24.00%	12.80%
14	域名在雅虎目录中被引用（链接本身以外的价值）	24.00%	12.20%
15	域名在开放目录中被引用（链接本身以外的价值）	23.00%	11.50%
16	域名在维基百科中被引用（链接本身以外的价值）	22.00%	12.40%
17	域名使用 Feeds	21.00%	10.80%
18	域名在互联网图书馆索引 lii.org 中被引用（链接本身以外的价值）	21.00%	12.40%
19	Google 网管工具中注册域名	18.00%	11.80%
20	开通 Google 增强型图片索引	17.00%	10.30%
21	域名使用安全证书（通过 https 处理交易）	14.00%	8.50%
22	域名注册信息中的地址、电话等记录有效	13.00%	8.30%
23	域名在 Google Knol 文章中被引用（链接本身以外的价值）	13.00%	9.20%
24	域名使用 Google 企业搜索工具（Google Search Appliance）	6.00%	7.40%
25	域名使用 Google　Adsense	5.00%	6.10%
26	域名使用 Google　Adwords 给自己做广告	5.00%	5.80%
27	域名的 Alexa 排名（与真实流量无关）	5.00%	5.80%
28	域名的 Compete.com 排名（与真实流量无关）	5.00%	6.10%
29	域名使用 Google 的企业应用套件（Google Apps）	3.00%	4.90%

全站非链接排名因素的评论如下。

Adam Audette：很多这些因素与 Google 怎样给一个域名排名并不直接相关，但是都对网站的 SEO 有很大影响，所以很难把它们一个一个拿出来看。我相信开放目录还很有

链接价值，Google 仍然使用开放目录就是个暗示。搜索一下“clothing”（衣服）这个词，看看前十名里总出现的两个结果，就可以得到更多暗示。这很有意义，尤其是没办法在开放目录被排名的分类页面上得到一个链接时。在高质量的网站目录中被引用，肯定是被跟踪和计算的，同时使用 Google 的域名检测手段。XML 网站地图对爬行有帮助，但本身并不是一个排名因素。

Marshall Simmonds：由于网站大小及内容深度原因，搜索引擎或者不在乎，或者没能力，或者不擅长对大型网站（几百万页的网站）进行自然爬行。也就是说，使用 XML 网站地图对企业级网站的成功来说至关重要，但小网站不一定能看到同样的好处。

Wil Reynolds：考虑到 Google 分析及 Google 工具条的广泛应用，Alexa 排名和 Compete 排名应该价值很低，Google 能从自己内部找到更多更准确的数据。

Richard Baxter：域名注册所有人的近期变化，尤其是如果域名曾经过期，对搜索结果有很负面的影响。

Ian Lurie：使用 Google Adsense、Google Apps、Google 搜索及其他搜索引擎提供的工具，完全不会影响排名结果。从 SEO 角度说，如果你的网站非得用 Google Adwords 广告才能被爬行抓取，你就有更大的问题了。

13.1.6 社会化媒体排名因素

社会化媒体排名因素如表 13-11 所示。

表 13-11 社会化媒体排名因素

排　名	因素说明	重要度	争议度
1	域名或网页的 Delicious 数据	21.00%	11.90%
2	域名或网页的 StumbleUpon 数据	19.00%	12.30%
3	域名或网页的 Twitter 数据	17.00%	10.70%
4	域名或网页的 LinkedIn 数据	15.00%	11.00%
5	域名或网页的 Face book 数据	12.00%	9.10%
6	域名或网页的 MySpace 数据	11.00%	8.40%

社会化媒体排名因素的评论如下。

Marty Weintraub：Twitter 数据还不是一个排名因素，但以后可能是。

Hamlet Batista：Matt Cutts 在一个视频里曾经说，Google 不在乎你有多少 Twitter 的 follower，他们的算法只在乎链接。

Dan Thies：这些不可能是排名因素，也永远不会是。

Todd Malicoat：社会化书签是一个质量指标，品牌名被提到是一个质量指标。如果我是搜索引擎工程师，我很可能会根据第三方社会化媒体上容易获得的有效数据，来对品牌搜索词进行排名。

13.1.7 用户数据排名因素

用户数据排名因素如表 13-12 所示。

表 13-12 用户数据排名因素

排 名	因 素 说 明	重 要 度	争 议 度
1	搜索结果中特定页面的历史点击率	42.00%	11.40%
2	搜索结果中一个域名所有网页的历史点击率	39.00%	11.30%
3	域名或品牌名称被当做查询词搜索	36.00%	12.30%
4	用户点击搜索结果后又调整搜索词	32.00%	11.20%
5	平均网页浏览时间	26.00%	12.00%
6	来自 Google SearchWiki 投票、评级、评论的数据	19.00%	9.10%
7	Gmail 邮件中出现的域名引用或链接	9.00%	7.70%

用户数据排名因素的评论如下。

Jessica Bowman：可用性可能是计算排名公式里的因素，但是我还没看到这会影响排名，尤其是对大型、有权重的网站来说。网站需要在这些方面投注精力，因为以后可能有很大影响。

Andy Beal：Google 可能正在实验在算法中包括这些因素，但我没看到这些因素被广泛使用的证据。

Adam Audette：点击率是一个很大的累计因素，页面调入时间也是如此，我们目前在这方面很关注。

Carlos Del Rio：自从 Vince 更新后，搜索词中加入品牌或域名，变得更重要了。

Ian Lurie：这些因素都还没有明显影响，但是就快了。如果你认为 Google 会忽视所有工具条数据和 SearchWiki 信息，你头脑就有点问题了。

13.1.8 负面排名因素

负面排名因素如表 13-13 所示。

表 13-13　负面排名因素

排　　名	因 素 说 明	重 要 度	争 议 度
1	恶意操控为目的的隐藏页面	68.00%	10.70%
2	从有名的链接买卖提供商那里得到的链接	56.00%	13.10%
3	页面上有连向垃圾网站或页面的链接	51.00%	12.10%
4	检测用户代理类型（浏览器还是蜘蛛），然后隐藏页面	51.00%	15.20%
5	服务器经常死机，网站不能打开	51.00%	12.30%
6	与背景相同或相似颜色的隐藏文字	49.00%	15.30%
7	域名上有连向垃圾网站或页面的链接	48.00%	13.10%
8	存在大量重复同样锚文字的外部链接到网站或网页	46.00%	11.00%
9	检测访问 IP 地址，然后隐藏页面	46.00%	15.30%
10	使用 CSS 隐藏文字，将文字置于可见区域之外	44.00%	14.80%
11	URL 中有大量动态参数	43.00%	13.50%
12	大量链接来自同样 C 段 IP 地址	41.00%	10.50%
13	意图操控的诱饵替换链接（将微型网站 301 转向等）	41.00%	12.90%
14	页面文字关键词堆积	41.00%	11.30%
15	使用 CSS display：none；隐藏文字	40.00%	14.20%
16	标题标签中关键词堆积	39.00%	11.20%
17	URL 中关键词堆积	37.00%	9.90%
18	通过操控意图的小工具或奖励、徽章活动获得的链接	37.00%	12.80%
19	检测 JS、富媒体支持，然后隐藏页面	37.00%	15.40%
20	检测是否支持 Cookie，然后隐藏页面	36.00%	16.30%
21	从低质量付费网站目录获得的链接	36.00%	12.20%
22	由同一个域名所有人的不同网站获得的大量链接	36.00%	12.40%
23	页面有来自垃圾网站和页面的链接	36.00%	13.10%
24	域名有来自垃圾网站和页面的链接	34.00%	14.00%
25	从意图操控排名的口碑传播活动获得的链接	33.00%	12.90%
26	以改进用户体验为目的的隐藏页面	33.00%	12.80%
27	内部链接锚文字过度优化	32.00%	11.20%
28	外部链接锚文字中包含某些黑名单关键词（例如学生信用卡、买伟哥、色情词汇等）	32.00%	11.90%
29	通过购买老域名然后转向获得的链接	32.00%	13.20%
30	超长的 URL	30.00%	13.00%
31	页脚使用关键词作为锚文字的内部链接	27.00%	10.20%
32	描述标签中关键词堆积	26.00%	11.20%

续表

排　名	因 素 说 明	重 要 度	争 议 度
33	购买老域名然后加上链接	24.00%	10.20%
34	内部链接过度使用 NF，控制 PR 流动	24.00%	10.90%
35	在论坛进行链接建设（签名、帖子里放链接）	22.00%	12.80%
36	超长的标题标签	21.00%	9.10%
37	关键词标签中关键词堆积	15.00%	10.90%

负面排名因素的评论如下。

Andy Beard：外部链接大量重复使用相同的锚文字：

- 这取决于这些链接是怎样为了长远利益获得的。
- 如果你设计一个 WordPress 模板，页脚上放上“买伟哥”链接，那就别寄望于人工审查人员会喜欢。

使用 CSS `display：none`；

隐藏文字：

- 这是不是导航系统的一部分，用户实际上能看到显示的内容。
- 如果你在这里隐藏一堆关键词或充满关键词的链接，那么可能就是一个很重要的因素了。

内部链接锚文字优化过度：如果是一个完美优化的链接，连向一个完美的登录页面，Google 不会因为这种你本来就应该做的事而惩罚你。链接实话实说就行了。

页脚使用含关键词锚文字的内部链接：

- 使用 CSS，你可以把页脚调到页头，或者把页头调到页脚。
- 在页面的这些地方放 100 个链接，对用户有意义吗？

购买老域名然后转向获得的链接：如果原来的老内容转向并放在新的域名上，这就不会有什么问题。

Debra Mastaler：你听到的很多关于小工具、微型网站、301 转向、购买老域名等会给你带来负面影响的说法，其实是一些人散布的夸张的恐怖战术的结果。这些方法有很多合理合法的用途，使用得当，作为一个整体市场计划的一部分是可以很成功的。

Tom Critchlow：很多这些因素取决于意图。比如根据用户代理（浏览器类型）隐藏网页，只要目的单纯就没问题。很多大型网站这么做很多年了也没有问题。一些链接因素，比如意图操控排名的诱饵替换手法，通常是没有价值的，而不是有负面影响。我们

发现 Google 在可能的情况下更喜欢对这些带有垃圾性质的策略/链接不给权重，而不是给予惩罚。

Carlo Del Rio：我还没见过购买老域名带来负面影响的，但通常对排名也没什么正面影响。目前来说，意图操控排名的链接是导致负面结果的最大威胁。重复锚文字、交叉链接、快速获得链接，就好像玩火一样，你早晚会被烧到。

Peter Meyers：链接农场的负面影响似乎是非常局限于特定利基主题的。有时 Google 非常严厉地打击，比如房地产行业。但是在一些小的利基市场，我看到一些人做链接农场也没什么事。我不太确认这种惩罚是否是在核心算法当中的。

13.1.9 影响外部链接价值的负面因素

影响外部链接价值的负面因素如表 13-14 所示。

表 13-14 影响外部链接价值的负面因素

排 名	因 素 说 明	重 要 度	争 议 度
1	域名因为垃圾被 Google 从索引中删除	70.00%	10.80%
2	域名排名因为垃圾被 Google 惩罚	65.00%	10.90%
3	链接被判定为付费的，而不是自愿给予的	63.00%	12.50%
4	域名链接到很多垃圾网站	52.00%	11.30%
5	域名还没有赢得被信任的链接	41.00%	11.80%

影响外部链接价值的负面因素的评论如下。

Adam Audette：都是重要因素，最后一个是灰色地带，但也是很重要的因素。如果一个链接被判定为是付费的，通常这个链接会被过滤出去。但是如果有太多付费链接，有时也可能导致严重惩罚。

Chris Bennet：我不知道 Google 用哪些标准通过算法找到低质量付费链接。但是建立一个工具，辨别出 80%～90%的垃圾应该不难。

Hamlet Batista：来自被删除网站的链接毫无用处。

Todd Malicoat：大部分链接并不会有害，但是如果你把精力都放在那些对你没什么帮助的链接上，你还是给自己带来了负面影响。应该把时间花在有用的链接上。

Ian McAnerin：链接不是一个排名因素，信任和主题才是。链接是信任和主题的代表。如果一个链接很少或基本没有信任度或者不太相关，那么这个链接就没什么价值。如果一个链接既没有信任度，也不能准确反映主题，就没有必要把它算进来。

13.1.10 地理定位因素

地理定位因素如表 13-15 所示。

表 13-15 地理定位因素

排 名	因 素 说 明	重 要 度	争 议 度
1	域名的国家代码（例如.co.uk .de .fr .com.au 等）	69.00%	7.90%
2	网站内容所使用的语言	63.00%	9.30%
3	外部链接所在域名针对的国家和地区	60.00%	10.30%
4	主机所在 IP 地址的地理位置	57.00%	12.00%
5	Google 工程师和质量评级人员人工检查和定位	53.00%	14.60%
6	Google 网管工具中设定的地理定位	52.00%	11.40%
7	注册 Google 本地商户	45.00%	10.30%
8	网页内容上的地址	41.00%	11.80%
9	域名注册信息中的联系地址	35.00%	12.30%
10	网站用户的地理位置（大部分网站用户来自于哪个国家地区）	30.00%	10.20%
11	通过 Meta 数据做的页面地理标签(例如都柏林核心元数据计划）	24.00%	10.80%

地理定位因素的评论如下。

Joost de Valk：在不同国家排名有不同要求，在某些国家 Google 并不能依靠 IP 地址可靠判定服务器地理位置。有些语言对 Google 算法来说显得很相近，所以有时会发生奇怪的事情（诸如荷兰文页面在德文搜索结果中出现）。

Russell Jones：有任何机会能够明确告诉 Google 你的网站是为哪个地区设计的，都不要错过。让他们的工作越简单越好。

Wil Reynolds：域名注册信息中的地址应该不会有很大作用，不然对那些在一个国家注册域名，但网站有针对多个国家内容的网站有很大伤害。

Aidan Beanland：据我的经验，Google 主要还是依靠域名国家代码，主机 IP 地址和网管工具中的地区设置。其他因素的作用，比对其他搜索引擎的重要性要低。

网站语言是一个自动的地域过滤器，因为只有那个语言的查询词才会在相应国家网站中出现。不过这也导致一些混乱，比如一种语言在多个国家使用时，或者一个单词在多种语言中使用时。

Kristjan Mar Haukson：我们工作中接触的一些大公司网站域名是在某一个国家注册的，但网站其实是针对另外一个国家的，我们并没看到这有什么影响。

13.1.11 附加 SEO 数据

接受调查的专家们也被问了下列问题，以帮助为 SEO 领域的一些关键争论提供真知灼见。

问题：下面哪种描述最符合你对 Google 2009 年 2 月品牌/Vince 更新的意见和体验？

- 51%：这次算法更新影响到一些排名因素，刚好使得著名公共品牌在搜索结果页面中显得有优先权（也不是全局都如此）。
- 36%：Google 现在给予著名公共品牌网站高一点的优先权。
- 9%：Google 现在给予著名公共品牌网站高得多的优先权。
- 4%：Google 结果没有明显偏向著名公共品牌。

问题：Google 算法在评估子域名上的内容（不包括特殊情况，如 Blogspot，Wordpress 等）方面，下面哪种描述代表你的意见？

- 83%：子域名内容继承一部分与搜索词无关的主域名排名因素，但不是全部。子域名一定程度上是被当做独立的实体。
- 10%：子域名内容并不继承与搜索词无关的主域名（或其他子域名）排名因素。子域名基本上是被当做一个独立的实体。
- 7%：子域名内容完全或基本上完全继承与搜索词无关的主域名（或其他子域名）排名因素。子域名差不多与所在主域名相同。

问题：在多大程度上你相信 Google 搜索使用 Google Analytics 收集到的数据影响其搜索排名？

- 74%：Google Analytics 数据只是以聚合形式被使用，帮助模式鉴别和整体用户行为分析。
- 16%：Google Analytics 数据没有以任何形式被使用。
- 6%：Google Analytics 数据被使用在网站上，可能正面或负面影响网站排名。
- 4%：Google Analytics 数据被使用在网站上，但只能正面影响搜索排名（没有进行垃圾或惩罚分析）。

问题：下面哪种描述最反映你对 Google 处理 301 转向的想法和经验？

- 70%：只有在某些内容和垃圾分析算法达到标准后，301 转向才从一个 URL 向另一个传递大部分（但不是百分之百）与搜索词有关或无关的排名因素。
- 23%：301 转向总是从一个 URL 向另外一个传递大部分（但不是百分之百）搜索词有关或无关的排名因素。
- 7%：301 转向总是百分之百地从一个 URL 向另外一个 URL 传递搜索词有关或无关的排名因素。

问题：就你的想法和经验，虽然使用了 NF，维基百科里的链接是否直接正面影响 Google 排名？

- 68%：是的。但是这些引用并不是直接被当做链接，而是作为潜在质量、权重、信任度的象征。
- 26%：不是。维基链接看起来传递价值，是因为很多其他网站和页面抄袭或使用了维基上的内容而没有使用 NF。
- 6%：是的。这些链接被处理为就好像 NF 不存在一样。

问题：未来 5 年，作为排名算法的一部分，Google 将怎样处理链接？下面哪种描述最符合你的意见？

- 48%：随着来自用户使用数据、社会化媒体数据和其他来源的信号取代链接，链接重要性将降低，但还保持原有效果。
- 37%：链接继续保持作为 Google 排名算法的主要部分，但是在计算链接数量及哪些链接有效方面会有重大波动。
- 15%：链接将继续保持为 Google 排名算法的主要部分，基本上和过去 5 年一样。
- 0：链接将变得基本过时，就像 20 世纪 90 年代关键词堆积没落一样。

13.1.12 链接建设调查

除了调查排名因素，我们也问了不同链接建设战术的有效性。既然获得链接是 SEO 非常重要的一部分，而链接又非常难以获得，我们感觉 SEO 专家们从他们自身及客户项目中获得的经验，将会给这份文件提供更多实质性的内容。在链接建设战术后面，我们列出了影响外部链接价值的因素，帮助你判断一个链接建设方法的回报或 ROI。

SEO 链接建设战术有效性，如表 13-16 所示。

表 13-16 SEO 链接建设战术

排 名	战术方法说明	重 要 度	争 议 度
1	链接诱饵+口碑传播内容创建	67.00%	8.80%
2	写博客并参与博客圈的讨论	66.00%	8.60%
3	传统的、创造有价值内容策略，但不积极推广	58.00%	12.40%
4	公共关系（除了发布新闻稿以外）	56.00%	11.70%
5	直接从单独的网站或站长那里买链接	54.00%	14.20%
6	小工具和其他可嵌入式内容	54.00%	11.10%
7	会议、活动或其他人际交往	54.00%	11.90%
8	用户产生内容（然后激励用户链接到个人介绍和内容页面）	53.00%	9.90%
9	高信任度和权威度的目录（开放目录、雅虎等）	52.00%	10.70%

续表

排　名	战术方法说明	重 要 度	争 议 度
10	利基社会媒体社区	51.00%	11.30%
11	本地链接建设（通过地域性列表、行业组织、门户等）	51.00%	10.10%
12	社交投票门户（如 Digg 、Reddit、Mixx 等）	50.00%	10.40%
13	创建问答题，然后颁发成绩徽章	50.00%	11.10%
14	社会化书签服务（如 StumbleUpon、Delicious 等）	49.00%	11.20%
15	向公益组织、非营利组织、活动组织等捐助，获得链接	45.00%	11.60%
16	充分利用 Twitter 建设链接	43.00%	11.80%
17	普通网站目录链接（如 BOTW、JoeAnt、Business.com 等）	42.00%	9.60%
18	直接联系站长或网站，请求获得非付费的链接	41.00%	14.90%
19	线下广告品牌建设以及媒体	39.00%	11.50%
20	新闻发布	39.00%	11.00%
21	长尾网站目录链接（如利基目录，小型普通目录等）	39.00%	11.50%
22	社交网络服务（如 Face book 、MySpace、LinkedIn 等	37.00%	8.50%
23	从链接买卖交易商那里购买链接	37.00%	12.90%
24	建立微型网站，然后 301 转向到主网站	36.00%	11.90%
25	购买老域名，然后在上面放链接	36.00%	14.50%
26	购买老域名，然后 301 转向到主站	32.00%	12.80%
27	互惠链接（与其他网站交换链接）	29.00%	8.20%
28	不带 NoFollow 的博客评论	29.00%	9.60%
29	网站广告（旗帜广告，PPC 等）	25.00%	10.50%
30	论坛链接建设（签名/帖子中的链接等）	23.00%	8.80%
31	自动博客、留言本和其他开放平台的评论垃圾	10.00%	8.00%

SEO 链接建设战术有效性的评论如下。

Jessica Bowman：公共关系（除了发布新闻稿以外）之所以没有发挥出它应该有的效率，原因之一是公关部门或代理速度不够快，他们的目的也不是把搜索引擎排名机会最大化。

Adam Audette：我们喜欢的一些策略：

- 创建一些非常高质量的内容，然后推广到我们设置的一些 Twitter 账号，把内容送到相应利基市场的博客写手们那里。
- 在 StumbleUpon 和 Reddit 购买付费广告，使某些内容初步变得热门，然后直接联系那些投票的人或者收藏的人，请求他们写客座博客或帖子。
- 联系站长、博客等，请求他们链接到我们网站上有价值的资源。

Wil Reynolds：我想利基目录和小型通用目录是两种不同的链接。利基目录能给你主题性的权威度，Google 根据链接关系能够判断与主题相关，所以比那些包含各种主题的小型通用目录更有价值。

Marcus Tandler：买老域名然后 301 转向到主站，有域名数量限制。一个域名很好，10 个域名 301 转向就不太好了。同样重要的是，这个域名应该是同样利基主题的。更重要的是，这些指向转向域名的锚文字中使用的关键词，在主站域名中至少要出现一次（比如说指向老域名的 100 个链接中包含锚文字“SEOmoz”，那么在转向的目标域名网站上，也应该包含关键词“SEOmoz”）。

Todd Malicoat：有很多链接建设的技巧。了解怎样评估一个链接的价值，怎样评估链接构成，就能了解哪种链接获得策略对你的网站排名及收入最有意义。

影响一个外部链接价值的因素如表 13-17 所示。

表 13-17　影响外部链接价值的因素

排　　名	因 素 说 明	重 要 度	争 议 度
1	源域名信任度（基于与被信任的种子网站链接距离的迭代计算，也就是 TrustRank）	70.00%	8.00%
2	源域名整体权威度/重要性（基于整站链接的迭代计算）	68.00%	9.00%
3	链接锚文字使用关键词（与搜索词匹配）	67.00%	8.40%
4	链接传递的 PR 值数量（例如页面具备的可传递的 PR 值除以页面上的链接数）	59.00%	10.40%
5	页面正文内容中的链接位置（不是侧导航、页脚等）	53.00%	10.50%
6	链接来源页面与目标页面的主题相关性	53.00%	10.30%
7	页面上链接的位置与其他链接的关系（周围有很多其他链接？还是包含在没有其他链接的正文内容中？）	46.00%	10.00%
8	链接来源域名与目标页面的主题相关性	46.00%	10.60%
9	页面上其他导出外部链接的质量	42.00%	11.60%

影响外部链接价值因素的评论如下。

Jon Myers：信任和质量是链接最重要的。确保你获得链接的页面有比较高的 PR 和信任度，你就成功了。找到链接之后的关键点还有锚文字集中、相关度及正确的目标页面。

Russell Jones：博客软文作为一个链接建设技术很成功，显示了包含在正文内容中的链接是有价值的。不过如果以为这是由于话题相关就不正确了。如果是这样，自动发布和同步文章就变得可行了。更可能的原因是，在独特内容中包含的链接被认为是真正自发自愿的链接。

Adam Audette：与目标页面匹配的内容中包含的链接非常有效。锚文字比链接周围的文字和页面关键词密度（换一种说法也就是页面的主题，这个页面是关于什么的）重要性小。侧栏导航中的链接和页脚中的链接很容易被 Google 判断出来并过滤掉。不过这些链接如果运用得当，是强有力的广告机会，甚至 NF 的链接也有可能带来大量流量，从而带来潜在链接。

Roger Monti：来自被信任网站的链接更能影响排名。链接的位置很重要。一个理想的链接是来自于主题相关、高度被信任的网页，位于网页顶部，包含在 H 标签或比较大的字体中。

Lisa D Myers：在页面正文中的链接比页脚或侧栏中的链接要好得多。在确定一个链接应该给予的相关性和能量时，链接周围内容相关性、语义分析索引都被大量使用。来自于一个被信任、内容相关网站的链接，永远有更高的能量。

13.2 Google 排名因素调查 2011

SEOmoz 的 Google 排名因素调查 2011 年继续进行。本节是调查结果的主要内容。

2011 年的调查采取了与 2009 年稍有不同的计算方法，没有按“重要度”和“争议度”表示，而是只计算出某因素从 0 到 100 的“影响值”。影响值越高，这个因素越重要。

排名因素按属性分为 9 类。每类因素列在不同表格中，其中列出的影响值是在本类内的相对值。

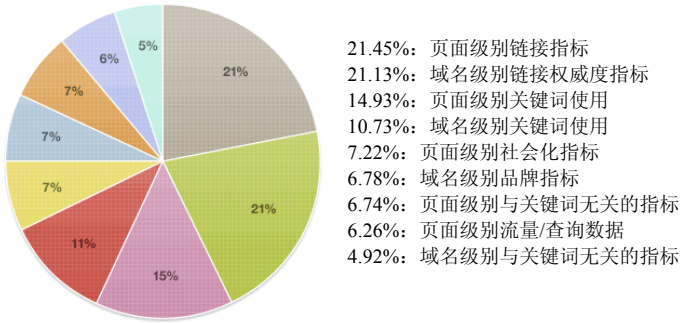


图 13-2　2011 年 Google 排名因素调查比重

13.2.1 域名级别关键词使用

域名或子域名中关键词是怎样使用的，以及对搜索引擎排名的影响，如表 13-18 所示。

表 13-18　域名级别关键词使用

排名因素说明	影 响 值
域名与关键词完全匹配（例如 keyword.com）	97.0
域名中出现关键词（例如 ABCkeyword.com）	71.0
关键词是域名中的第一个词（如 keywordABC.com）	69.3
子域名是关键词（如 keyword.ABC.com）	52.8
次级子域名是关键词（如 keyword.123.ABC.com）	27.2
域名扩展名是关键词（如 ABC.keyword）	17.2
域名中的字母组合形成关键词（如 key.word.com 或 hey.wo.rd）	15.2

13.2.2　域名级别链接权威度指标

页面所在域名的链接指标，如表 13-19 所示。

表 13-19　页面所在域名的链接指标

排名因素说明	影 响 值
域名获得的独特链接域名总数	80.4
域名的“原始”重要度，以域名获得的所有链接重要度总和衡量（如 mozRank 和 Google PR 等计算的）	67.3
以 TrustRank、mozTrust 等算法计算的域名与被信任种子网站的“距离”	57.4
域名获得的所有独特链接 URL 总数	55.5
链接域名的主题相关性	52.5
链接获得的加速度（如大量链接快速出现，指向网站的不同页面）	47.7
链接向域名的独特 c 段 IP 地址总数	46.6
锚文字包含品牌名称的链接百分比	46.0
连向域名的链接类型多样性（如论坛、博客、新闻）	44.6
来自被 Hilltop 之类算法认为是枢纽页面的链接	43.7
指向网站的外部链接的评价（译注：指含有正面、负面文字意义的评价）	8.3

13.2.3　域名级别与关键词无关的指标

与整个域名有关但不是描述链接或关键词的元素。与域名长度、网站错误页面总数、网站内容相对独特性之类的有关特征，如表 13-20 所示。

表 13-20　域名级别与关键词无关的指标

排名因素说明	影 响 值
整个网站内容独特性	89.4

续表

排名因素说明	影 响 值
网站内容更新度	74.9
网站弹出率，以用户访问网站后很快返回搜索结果页面衡量	65.2
域名在 Google 搜索结果页面的总体点击率	62.2
网站被爬行的错误页面总数	60.5
域名第一次被注册的时间	52.8
域名上的页面总体 HTTP 反应时间（Google 蜘蛛测量的）	49.3
域名上页面总体打开速度（用户数据测量的）	47.1
域名链接向多少 404 页面	45.0
域名注册信息保持不变有多久	44.9
同一段 c 段 IP 地址上其它网站的质量	31.5
域名中使用了连线符	30.2
域名有多少个字符	28.3
域名还有多久过期	18.5

13.2.4 域名级别品牌指标

描述域名品牌质量的元素，如表 13-21 所示。

表 13-21 域名级别品牌指标

排名因素说明	影 响 值
品牌名/域名搜索量	84.7
品牌名/域名在新闻网站被提到的次数	78.0
品牌名/域名在社会化网站被提到的次数	68.8
工具条/浏览器记录的域名使用数据	60.6
域名在 Wikipedia 被引用	60.4
企业已在 Google Places 注册（译注：谷歌本地或谷歌地图）	55.0
网站有官方 Facebook 页面	34.3
网站有官方 Twitter 页面	33.5
网站有官方 Linkedin 页面	28.7
网站经常在 Linkedin 用户信息中被列为雇主	25.1
与网站有关的企业有交税记录	20.9

13.2.5 页面级别社会化指标

与第三方社会化媒体资源（Facebook、Twitter 等）有关的指标，如表 13-22 所示。

表 13-22　页面级别社会化指标

排名因素说明	影 响 值
Twitter 上推（tweet）页面链接的用户的权威度	84.1
页面被推（tweet）的次数	79.0
页面在 Facebook 被分享（share）的次数	61.6
Facebook 上分享（share）页面链接的用户的权威度	56.3
页面在社会化网站（如 Digg、Reddit、Stumbleupon、Mixx 等）被投票	40.9
Google Buzz 上分享（share）页面链接的用户的权威度（译注：Google Buzz 已停止服务）	36.9
页面在社会化网站（如 Digg、Reddit、Stumbleupon、Mixx 等）上的评论	35.6
页面链接在 Google Buzz 上出现的次数	34.4
社会化链接和引用对页面的评价	21.3

13.2.6　页面级别链接指标

单独页面的链接指标（如链接数目，mozRank 等），如表 13-23 所示。

表 13-23　页面级别链接指标

排名因素说明	影 响 值
以关键词为锚文字的独特链接域名总数	71.9
以关键词为锚文字一部分的独特链接域名总数	71.5
指向页面的独特链接域名总数	69.1
关键词与锚文字完全匹配的链接总数	63.8
关键词是锚文字一部分的链接总数	60.1
以 PR、StaticRank、mozRank 等算法计算的页面“原始”重要度	55.7
链接页面的主题相关性	51.6
以 TrustRank，mozTrust 等算法计算的页面与被信任种子网站的“距离”	50.5
链接域名的主题相关性	42.9
链接向页面的独特 c 段 IP 地址总数	40.9
链接向页面的独特 URL 总数	40.1
连向页面的链接类型多样性（如博客、论坛、目录）	38.3
来自被 Hilltop 之类算法认为是枢纽页面的链接	36.4
指向页面的外部链接的评价	7.2

13.2.7　页面级别关键词使用

描述页面 HTML 代码的特定部分（标题，H1，图片 ALT 属性等）使用关键词情况，如表 13-24 所示。

表 13-24　页面级别关键词使用

排名因素说明	影 响 值
标题标签中出现关键词	94.4
标题标签的第一个词是关键词	77.8
指向其他内部页面的链接锚文字中出现关键词	72.9
页面 URL 中出现关键词	69.9
页面 H1 标签中出现关键词	69.7
指向外部页面的链接锚文字中出现关键词	67.2
页面“内容”部分（搜索引擎通过对导航/边导航条/页脚/内容进行某种视觉分区计算决定）出现主题关键词	65.9
文件首 100 个字中出现关键词	63.7
根据主题模型算法（如 LSA/LSI/pLSI/LDA 等）得到的页面整个内容相对用户查询的优化程度	60.4
页面图片 ALT 属性中出现关键词	51.1
页面 H2 标签中出现关键词	47.2
文件可见文字第一个词是关键词	38.7
关键词是文件里相同长度的词（一个字的词，两个字的词等）中最常出现的	38.5
页面上图片的文件名中使用关键词	37.7
黑体/斜体标签中出现关键词	37.2
文件中关键词出现 3～6 次（绝对数量）	37.1
关键词密度（页面上目标关键词的百分比）是“完美”值	33.0
页面描述标签中出现关键词	31.9
页面 H3 标签中出现关键词	29.3
描述标签的第一个词是关键词	21.7
页面关键词标签中出现关键词	5.2

13.2.8　页面级别流量数据

页面的用户及使用数据（假设 Google、Bing 通过它们的工具条、浏览器、移动设备可以获得大量流量数据用以分析），如表 13-25 所示。

表 13-25　页面级别流量数据

排名因素说明	影 响 值
页面在 Google 某个关键词结果中的点击率	74.6
通过返回搜索结果页面的访问得到的页面弹出率	58.4

续表

排名因素说明	影 响 值
页面在 Google 所有关键词的总体点击率	51.4
页面还获得排名的其他关键词数量	41.4
页面从 Google 获得的流量	37.6
浏览器/工具条收集的页面使用数据	36.6

13.2.9 页面级别与关键词无关的指标

页面与关键词及链接无关的元素（如页面长度、打开速度等），如表 13-26 所示。

表 13-26 页面级别与关键词无关的指标

排名因素说明	影 响 值
页面内容独特性	93.7
域名上是否有其他页面链接到这个页面（也就是页面是否是个“孤岛”）	82.7
页面内容是否新鲜（也就是页面最后一次更新的时间）	70.5
页面年龄（也就是页面第一次被发现的时间）	63.8
页面内容长度	57.4
蜘蛛测量的页面 HTTP 反应时间	53.3
用户工具条测量的页面打开速度	47.1
页面上的图片使用	46.9
页面上是否有链接指向 404 页面	44.9
页面上微格式（microformat）的使用（在相关垂直领域）	44.2
页面上视频的使用	33.3
页面上广告的使用	32.1
页面内容的阅读水平	31.0
页面上指向外部的联署计划链接使用	30.3
W3C 标准 HTTP 验证	18.9

13.3 百度排名因素调查 2010

这个百度排名因素调查是我专门为本书而做的。仿照 SEOmoz 排名因素调查的方式，向一些我认识的国内 SEO 专家和同行发出问卷，共收回 26 份有效回复。

下面的排名因素都给出 5 个评分选择。

- 1 分：非常不重要。
- 2 分：不太重要。

- 3 分：一般。
- 4 分：比较重要。
- 5 分：非常重要。

由被调查专家评分，表中列出的重要度是所有专家评分的平均值。

调查前言中，我强调了几点：

- 对每一个问题，请按照对百度（仅针对百度，请不要考虑 Google 排名）的个人经验和判断评分。
- 请不要参考 SEOMoz 的调查，也不要参考他人意见。
- 调查中的问题是随机排列的，先问的因素不意味着重要性比较高，请不要受问题顺序的影响（调查问卷中的问题没有按重要性排序）。

感谢参与的和由于时间及技术原因最终没能参与的同行们，相信这份百度排名因素调查对中文 SEO 所有从业人员和爱好者都有很大的借鉴意义。

13.3.1 与关键词有关的页面排名因素

与关键词有关的页面排名因素如表 13-27 所示。

表 13-27 与关键词有关的页面排名因素

排　名	排名因素说明	重 要 度
1	标题（title）标签中任何地方使用关键词	4.65
2	关键词完整匹配出现在页面，关键词邻近度 如目标关键词“北京旅游”，“北京旅游”出现在页面比“北京秋季旅游”有效	4.15
3	标题标签中第一个字或词使用关键词	4.04
4	H1 标签中任何地方使用关键词	3.81
5	页面最前面 50～100 个可见文字中出现关键词	3.81
6	关键词密度 关键词使用次数除以页面上总词数	3.81
7	页面可见文字中关键词使用和重复次数 是否关键词重复次数越多排名越好	3.69
8	H1 标签中第一个字或词使用关键词	3.62
9	标题标签（title tag）中适当重复关键词 如 2～3 次	3.58
10	页面上导出链接锚文字中使用关键词	3.43
11	图片 ALT 文字中使用关键词	3.38

续表

排　　名	排名因素说明	重 要 度
12	描述标签（description tag）中使用关键词	3.38
13	黑体（bold 或 strong）中使用关键词	3.35
14	其他正文标题（H2～H6）中使用关键词	3.23
15	页面上列表（li）中出现关键词	3.04
16	页面最后面 50～100 个可见文字中出现关键词	2.85
17	目录名中包含关键词中文拼音 如目标关键词为“北京旅游”，目录名为 /beijing-lvyou/	2.81
18	文件名中包含关键词中文拼音 如目标关键词为“北京旅游”，文件名为 beijing-lvyou.html	2.77
19	页面上使用的图片文件名中包含关键词 无论是拼音，对应英文或中文，如图片文件名为 beijing-lvyou.jpg	2.73
20	关键词标签（keywords tag）中使用关键词	2.73
21	子域名中出现关键词中文拼音 如目标关键词为“北京旅游”，子域名为 beijing-lvyou.domain.cn	2.65
22	斜体（italic 或 em）中使用关键词	2.65
23	域名中包含关键词中文拼音 如目标关键词为“北京旅游”，域名为 beijing-lvyou.cn	2.62
24	目录名中包含中文关键词 如目标关键词为“北京旅游”，目录名为 /北京旅游/	2.58
25	文件名中包含中文关键词 如目标关键词为“北京旅游”，文件名为 北京旅游.html	2.5
26	目录名中包含关键词对应的英文词 如目标关键词为“北京旅游”，目录名为 /beijing-travel/	2.5
27	子域名中包含关键词对应的英文词 如目标关键词为“北京旅游”，子域名为 beijing-travel.domain.cn	2.42
28	文件名中包含关键词对应的英文词 如目标关键词为“北京旅游”，文件名为 beijing-travel.html	2.38
29	域名中包含关键词对应的英文词 如目标关键词为“北京旅游”，域名为 beijing-travel.cn	2.27
30	HTML 注释标签（<!-- 注释文字 -->）中使用关键词	1.65

13.3.2　与关键词无关的页面排名因素

与关键词无关的页面排名因素如表 13-28 所示。

表 13-28　与关键词无关的页面排名因素

排　　名	排名因素说明	重 要 度
1	首页优势 即使链接情况一样，网站首页比内页排名有优势	4.42
2	独特的原创内容	4.31
3	URL 静态化	3.88
4	页面更新的新鲜度 页面最后一次更新的时间，是否刚刚更新的页面有优势	3.77
5	HTML 中格式代码与可见文字之比 代码精简是否有助于排名	3.77
6	页面 URL 目录深度 即使链接情况一样，一级目录下的页面比二级目录下页面排名有优势	3.65
7	页面上连到同一个域名下的其他 URL 导出的内部链接是否有助于本页排名	3.42
8	页面内容历史变化 页面内容更新频繁是否有助于排名	3.38
9	页面创建的时间 是否页面越老，排名越好	3.35
10	使用描述标签	3.35
11	W3C 标准验证	2.96
12	页面上指向其他网站的链接 导出的外部链接是否有助于排名	2.62
13	页面上有第三方展示广告 有展示广告是否说明网站比较重要	2.23
14	使用 Flash 或其他需要插件才能显示的内容	1.88

13.3.3　特定页面链接流行度排名因素

特定页面链接流行度排名因素如表 13-29 所示。

表 13-29　特定页面链接流行度排名因素

排　　名	排名因素说明	重 要 度
1	外部链接锚文字使用关键词	4.77
2	外部链接流行度 除整个网站外，连到本页面的外部链接数量及质量也影响排名	4.69
3	特定页面的 TrustRank 这个页面是否从被信任的网站获得了链接	4.54

续表

排　　名	排名因素说明	重 要 度
4	链接来源多元化 本页面的外部链接来自很多独特域名	4.42
5	导入内部链接锚文字使用关键词	4.31
6	外部链接来源内容相关性 页面的外部链接是否来自于主题相关的页面和网站	4.12
7	内部链接流行度 同一个域名下其他页面连过来的反向链接	3.96
8	页面在网站结构中的位置 页面在网站结构关系图中处于什么样的位置？如栏目首页比内容页面有优势	3.88
9	基于迭代算法得到的整体链接流行度 如本页面的 Google PR 值	3.73
10	导入 Nofollow 链接的数量质量 百度是否支持 Nofollow	2.31
11	页面的导入 Follow 和 Nofollow 链接比例	2.23

13.3.4　全站链接有关排名因素

全站链接有关排名因素如表 13-30 所示。

表 13-30　全站链接有关排名因素

排　　名	排名因素说明	重 要 度
1	来自特定主题的权威网站的链接 类似 Hilltop 算法	4.15
2	域名链接多元化 指向这个域名所有链接的域名总数量	4.15
3	基于与被信任域名链接距离的域名信任度 类似 TrustRank 算法	4
4	来自不能随便注册的域名的链接 如来自.edu，.gov，.mil 等域名的链接是否权重更高	3.88
5	指向域名的链接短时间内快速增长或减少 一段时间内获得外部链接的数量质量急剧增长是否使排名很快上升，急剧下降是否有负面影响	3.73
6	基于迭代算法得到的整体链接流行度 如域名 PR 值	3.5
7	指向域名的 Follow 和 Nofollow 链接比例	2.46

13.3.5 全站非链接相关排名因素

全站非链接相关排名因素如表 13-31 所示。

表 13-31 全站非链接相关排名因素

排　名	排名因素说明	重 要 度
1	域名被 hao123 和 265 等著名网址站收录	4.42
2	网站架构是否清晰有效	4.23
3	服务器在线率和稳定性	4.19
4	域名是否包括在百度新闻源中	3.85
5	域名注册历史 域名已经注册存在并归属同一个人多久？老域名有优势吗	3.69
6	网站打开速度	3.65
7	域名被开放目录收录	3.62
8	使用百度推广 使用百度凤巢及竞价排名有助于自然排名	3.31
9	导出链接到其他被信任有权重的网站或网页	3.12
10	在维基百科中域名被引用 链接本身以外的价值	3.12
11	主机信息 同一台主机和 IP 地址上还有其他网站影响排名吗	3.08
12	域名被提到情况 域名和网址以文字形式在其他网站被提到，但不是链接	3.08
13	域名已续费多少年 续费比较久是否意味着站长更认真	2.96
14	网站使用 Feed 可以订阅的 RSS	2.73
15	在互联网图书馆索引 archive.org 中域名被引用 链接本身以外的价值	2.73
16	百度博客搜索中有这个域名的 Feeds	2.69
17	网站参加了百度联盟	2.58
18	使用 XML 网站地图 百度是否支持 XML 网站地图	2.54
19	网站使用百度搜索框	2.46
20	域名的 Alexa 排名	2.35

续表

排　　名	排名因素说明	重 要 度
21	网站使用百度统计	2.12
22	域名注册信息中的地址电话等记录真实有效	2
23	域名使用安全证书 通过 https 处理交易有助于权威性	1.96
24	域名所有人的改变 根据注册信息，域名是否曾经转手	1.92
25	域名所有人 谁注册了域名？域名所有人拥有其他域名影响本网站排名吗	1.65

13.3.6　社会化媒体排名因素

社会化媒体排名因素如表 13-32 所示。

表 13-32　社会化媒体排名因素

排　　名	排名因素说明	重 要 度
1	域名或网页出现在百度百科、贴吧、词典、知道	3.42
2	域名或网页的新闻聚合数据 Digg，StumbleUpon 等地方出现网站的推荐	2.42
3	域名或网页的社交网络数据 LinkedIn，Facebook，MySpace 等	2.35
4	域名或网页的微博数据 网站 URL 出现在 Twitter，新浪微博等	2.31
5	域名或网页的网络书签数据 Delicious（美味书签）收录次数等	2.19

13.3.7　用户数据排名因素

用户数据排名因素如表 13-33 所示。

表 13-33　用户数据排名因素

排　　名	排名因素说明	重 要 度
1	搜索结果中一个域名所有网页的历史点击率	3.69
2	用户直接搜索域名或相关品牌	3.69
3	搜索结果中特定页面的历史点击率 所谓点击器是否有效	3.58

续表

排　　名	排名因素说明	重 要 度
4	搜索结果被点击后用户调整搜索词 如搜索“SEO”后没找到点石网站，用户转而搜索“SEO 点石”对点石网站排名有帮助	3.54
5	平均网页浏览时间，跳出率，用户每次访问浏览页面数量等 表现用户体验满意度的数据是否影响排名	3.38

13.3.8　负面排名因素

说明：负面排名因素与其他因素相反，如表 13-34 所示。

非常不重要：意味着负面影响非常小，没关系，不必在意。

非常重要：被发现的话负面影响非常大，所以不建议使用。所以非常重要，反倒是不要使用。

表 13-34　负面排名因素

排　　名	排名因素说明	重 要 度
1	百度工程师人工调整	4.23
2	页面文字关键词堆积	4.15
3	标题标签中关键词堆积	4.15
4	服务器经常宕机，网站不稳定	4.12
5	频繁更改标题标签	4
6	导出链接到垃圾网站或网页	3.88
7	页面文字中包含某些黑名单关键词 例如私服、六合彩、色情词汇等	3.88
8	网站被判定为参与链接买卖	3.88
9	外部链接锚文字中包含某些黑名单关键词 例如私服、六合彩、色情词汇等	3.81
10	隐藏文字 与背景颜色相同/相近，CSS 隐藏，微小文字等	3.77
11	描述标签中关键词堆积	3.77
12	URL 中关键词堆积	3.73
13	隐藏页面（cloaking） 搜索引擎看到的页面内容和用户看到的不同	3.69
14	来自垃圾网站和页面的外部链接	3.69

续表

排　　名	排名因素说明	重 要 度
15	内部链接锚文字过度优化 使用太多关键词做锚文字	3.69
16	大量来自相同或邻近 IP 地址的外部链接	3.65
17	关键词标签中关键词堆积	3.65
18	URL 中存在大量动态参数	3.62
19	超长的 URL	3.62
20	URL 跳转 JavaScript，302，Meta refresh 等	3.58
21	超长的标题标签	3.5
22	从低质量付费网址站获得的链接	3.46
23	大量同样锚文字的外部链接	3.27
24	从同一个域名所有人的大量网站获得的大量链接	3.23
25	友情链接过多	3.19
26	大量博客垃圾评论链接	3.12
27	购买老域名然后做转向获得的链接 买老域名，然后做 301 转向到主网站，老域名原来的链接也成为主网站链接	2.96
28	从有名的链接买卖中间商那里得到的链接	2.92
29	页脚处大量关键词为锚文字的内部链接	2.92
30	购买老域名然后加上链接 保留老域名内容，加上链接到主网站	2.58
31	内部链接过度使用 Nofollow，意图控制链接权重流动 大量使用 Nofollow，是否会被认为刻意优化	2.58
32	在论坛进行链接建设 大量在签名、帖子里放链接	2.58
33	网站使用 Google Adsense Adsense 对百度排名有负面影响	2.23
34	网站没有使用百度推广	2.23

13.3.9　地理位置定位因素

对有关地理位置的搜索（如“北京律师”），不同地理特征的网站排名是否不同。用户所在地理位置不同，看到的搜索结果是否不同，如表 13-35 所示。

表 13-35　地理位置定位因素

排　　名	排名因素说明	重 要 度
1	外部链接源域名的地理定位 来自上海有关网站的外链多，有助于“上海”相关词或上海用户排名	3.27
2	网站用户的地理位置 大部分网站用户来自于北京，是否对所有北京用户排名都有提升	3.19
3	地区域名 例如.bj.cn 域名有助于“北京”相关关键词或来自北京用户的排名，.sh.cn 域名有助于“上海”相关词	3.04
4	主机所在 IP 地址的地理位置 主机在北京，有助于“北京”相关关键词或来自北京用户的排名	2.62
5	网页内容上的联系地址	2.35
6	域名注册信息中的联系地址	2.31

感谢下列 SEO 专家和同行参与百度排名因素调查，如表 13-36 所示（按回复调查的时间顺序）。

表 13-36　SEO 专家及同行

姓　　名	简　　介
oxygen 真名：黄海均	个人简介：网易，产品设计师。长期关注 Web 2.0、电子商务及社会化媒体营销，目前在网易做互联网产品设计及运营相关工作。 个人博客或网站：http://ecvip.org
David Yin 真名：David Yin	个人简介：网站优化推广的研究实践者。 个人博客或网站：http://seo.g2soft.net/
潜索王道 真名：马骏	个人简介：online marketing executive@globalsources，龙之潜，乃龙之必腾。 个人博客或网站：http://www.zaccode.com/247
穆图 真名：杜晨	个人简介：搜索引擎营销观察组织成员，《轻公司》一书作者，互联网创业者。 个人博客或网站：http://semwatch.org
独孤天骄	个人简介：SEO 高级工程师，SEO 研究院博主，雅虎 SEO 大赛一等奖获得者，多家 SEO 公司顾问。 个人博客或网站：http://www.dugutianjiao.com
Kyw 真名：康轶文	个人简介：上海易闻网络科技有限公司创始人。中国搜索营销行业后起之秀，为中国 SEO 划分等级。2007 年创办上海易闻搜索营销公司并担任 CEO。擅长以实验的方式研究搜索引擎营销及网站优化技术，愿意分享经验，写作风格通俗易懂。 个人博客或网站：http://www.SEOtest.cn
猎眼 真名：陈恭	个人简介：上海易点网络，职业 SEO。 个人博客或网站：http://www.zhuici.com/blog/post/category/lie-yan

续表

姓　名	简　介
Money 真名：谢永钱	个人简介：金蝶友商网网络营销助理，online marketing 新人。 个人博客或网站：http://semwatch.org
天真 真名：葛小飞	个人简介：Brothersoft.com，继续摸索 SEM 的人。
Robert@ SEM 真名：葛珲	个人简介：刚刚上路的网络营销人。
阿猎 真名：周睿	个人简介：Netconcepts 中国区首席 SEO 顾问。先后为中国的行业 B2B 门户，B2C 网站提供整站 SEO 优化、数据分析挖掘等工作，其中包括中国最大的 B2B 网站慧聪网。一直从事策划、运营、UE 等多方面研究。 个人博客或网站：http//www.etcis.com
shaowei 真名：邓少炜	个人简介：关注新媒体，搜索引擎营销和电子商务。 个人博客或网站：http://semwatch.org
sem123 真名：陈沿舟	个人简介：杭州思亿欧网络科技有限公司，从事搜索引擎优化服务行业。 个人博客或网站：http://www.seo.com.cn/
樂思蜀 真名：王志炜	个人简介：西安欧派信息技术有限公司经理。四川人，现居西安，SEO 论坛（SEOBBS.Net）站长，专业研究 SEO 与 SEM 技术，擅长于多种营销方式与搜索引擎的营销相结合，最大化 ROI，先后为上百家企业提供搜索引擎优化与营销服务。 个人博客或网站：http://www.lesishu.cn
姜东栋 真名：姜东栋	个人简介：上海奇商网络科技有限公司 CEO。上海 SEO 联盟发起人之一。 个人博客或网站：http://www.qisir.com
Charles 真名：曹艳华	个人简介：南京科泰 在线营销总监。专注英文网站推广。 个人博客或网站：http://www.kseo.cn
黑色梦中 真名：郝聪	个人简介：SEO 技术研究者。 个人博客或网站：http://www.bloghuman.com
thinkingit 真名：周源	个人简介：Meta 搜索 CEO，联合创始人。 个人博客或网站：http://blog.meta.cn
国宝 真名：李国宝	个人简介：President，Good boy。点石最资深版主之一。 个人博客或网站：http://www.pricebat.ca
hi3w 真名：王峰	个人简介：深圳市天艺网络技术有限公司英文 SEO 主管。专注于 SEO、网站运营的研究和实践，较丰富的英文 B2C 网站运营经验。 个人博客或网站：http://www.zaccode.com/1030
张何 真名：张何	个人简介：企赢网络营销策划机构创始人。1998 年开始关注互联网，在电子商务、行业网站推广、搜索引擎优化、网站策划运营、赢利模式分析方面具有独到的心得。 个人博客或网站：http://www.xzsem.com

续表

姓　　名	简　　介
酒醒 真名：陈炳柱	个人简介：酒醒还是我。
核桃 真名：何涛	个人简介：杭州志卓 SEO。成长中的 SEO 实践者。 个人博客或网站：http://www.abseo.cn/blog/
charlie.Qin 真名：秦桂东	个人简介：慧聪 SEO 工程师。热爱 SEO、尊敬 SEO。 个人博客或网站：http://www.hc360.com
小彭 真名：彭亚运	个人简介：SEO 优化专员。 个人博客或网站：http://www.pyy1990.cn
封伟军 真名：封伟军	个人简介：SEO 主管。2007 年入行一直到现在，一直使用白帽 SEO 手法，做 SEO 应该不断地去研究 SE 和讨好 SE。 个人博客或网站：http://www.seoxn.com

第 14 章
SEO 案例分析

市面上已经有不少 SEO 书了。在规划本书内容时，我一直在琢磨怎样才能让这本书与众不同，而且对想实践 SEO 的读者最有助益。研究了一些 SEO 书和教程后觉得，一个完整、详细的案例可能是个亮点。不是那种简单几页的分析说明，而是非常完整、详细的案例，从竞争研究到网站诊断，还包括优化细节说明。

选择案例时经过了一些思考。

首先，网站不能太大，也不能太小，不然对普通站长和企业的借鉴意义不大。毕竟，不可能人人做个上千万页面的门户，分析几十几百页的小网站又会漏掉很多重要内容。所以案例网站应该在几十万至几百万页面之间。最后选定的亿赐客网站是几百万页面规模。

网站不能是和自己无关的，不然所谓的案例就只能是从表面现象分析，而不一定能看到本质。经常在网上看到分析成名网站 SEO 方法的文章，但其实分析别人的网站往往是不可靠的，因为很多时候网站上存在的东西外人无法知道其真正原因。就算仔细研究一个 SEO 非常成功的网站，他们使用的方法可能是很多因素折中后的结果，而不是最好的方法，甚至可能是他们 SEO 部门深恶痛绝却没有权力改变的。当然，分析别人的网站也是学习 SEO 技巧的途径之一，但首先需要学会鉴别，以免被误导。

只有自己经手优化的网站才能从里到外了解，知道每一个地方改了什么，为什么改，改了以后有没有效果。所以选定的网站成为本书案例合作伙伴，我出报告，他们可以实施优化。

网站不怕被曝光。这也是不能选以前 SEO 服务客户的原因之一。很少有公司愿意透露网站数据和优化过程，因为要冒着被搜索引擎注意和被同行抄袭的风险。亿赐客团队对此的态度是：我们坚信自己的网站在本行业里用户体验最好、技术最优秀、保证完全使用白帽，不怕被曝光、被抄袭。

最重要的是，网站不能是以前的项目，不然说服力不大。谁都能拿一个已经成功的网站来说事，不成功的就不提了。所以本章中的案例是专门为本书选择和优化的，并且案例网站在本书第一版写完前就公布了。我的优化报告于 2009 年 10 月份提供给亿赐客，2010 年 1 月第一个优化版本上线，之后又经过多次的观察和修改。

所以，这个案例是个真正详细说明优化过程及效果的案例。我在写优化报告和本书其他章节时，自己都还不知道效果会是什么样。

我希望，这是这本书最有说服力的地方。我还没见过这样的案例。大家都明白，这是要冒风险的，除了曝光、抄袭的顾虑，毕竟搜索引擎不是我们开的，谁也不能保证什么。但我还是决心实施并公布这样一个案例。要做，就做不一样的。哪怕案例失败，相信很多人也能从中获取经验教训。

看过我博客的人一定都知道，在我开始写这本书时就确定了这样一个案例，而且间接提到过案例网站，所以，冒险公布未完成案例是一开始就决定的，而不是看到有效果后再来公布。

现在距离案例网站操作已经过了两年，一些情况有了变化，搜索引擎算法有改变，亿赐客网站也有了不少并非基于 SEO 考虑的变化。但是，SEO 的基本原理和技巧并没有变化，而且我预计接下来两三年内也不会有大的变化。所以，案例中的竞争对手分析、网站本身诊断过程和优化方法目前同样适用。

希望下面的案例详情对广大站长和企业有所帮助。

14.1 竞争对手分析

亿赐客（www.yicike.com）是个比较购物网站。除了产品说明，用户还可以在网站上看到产品在不同电子商务网站的报价和一些其他信息如促销、配送方式等，帮助用户选择最好的购买网站。

2009 年 6 月刚刚接触对方时，其名称为“亿枝客”。2010 年才改为“亿赐客”。

首先，确定主要竞争对手相对简单。在搜索引擎输入“比较购物”，再简单查看一些相关文章，就可以大致确定几个主要竞争对手网站：

- 聪明点
- 智购
- 大拿网
- 丫丫网
- 特价王

在开始这个项目之前，我对中文比较购物网站没什么了解。搜索以后确定的几个竞争网站，与从亿赐客那里得到的信息基本一致。

14.1.1 了解网站基本数据

尤其是外部链接和社会化网站数据。可以使用 SEOBook 的 SEO for Firefox 插件（本节提到和使用的插件和查询工具在本书 SEO 工具部分都有详细介绍），如图 14-1 所示。

图 14-1 说明，搜索“比较购物”时，图中几个网站并没有连续排在一起。为显示和比较方便，将几个主要竞争对手网站剪裁到一个图片中。这是本书唯一经过处理的抓图。

本章对各相关网站分析的图片绝大部分为 2009 年 7 月份的抓图。本书引用的竞争对手网站资料均来自网上免费、公开的数据，任何人都可以查询，没有任何商业机密。

如图 14-1 所示，各个竞争对手的基本情况一目了然。表 14-1 是从抓图中挑出的几个最重要数据。

图 14-1　2009 年 7 月 12 号几个比较购物网站数据抓图

表 14-1　2009 年 7 月 12 号记录的竞争对手数据

	Google PR	首页快照新鲜度	年　　龄	雅虎全站外链	雅虎首页外链	Google 收录	开放目录
聪明点	6	2009 年 7 月 10 号	2003 年 10 月	1 250 000	1 200 000	6 640 000	收录
智购	5	2009 年 7 月 11 号	2004 年 4 月	163 000	69 900	1 580 000	无
丫丫	6	2009 年 7 月 8 号	2004 年 12 月	154 000	130 000	471 000	无
特价王	5	2009 年 7 月 10 号	2001 年 2 月	220 000	72 700	289 000	收录
大拿网	6	2009 年 7 月 9 号	2005 年 10 月	426 000	330 000	842 000	收录

相比之下，亿赐客域名 2007 年 8 月才注册，网站首页 PR 值 5，快照日期没有显示，雅虎全站链接 2 万左右，Google 收录 6 万多。可以看到，亿赐客网站除了年龄最短这个无法改变的劣势外，外部链接数量不到最弱竞争对手的15%。增加外部链接，提高网站权重，进而改善收录和长尾关键词排名将会是个艰巨的任务。

14.1.2 外部链接

从外部链接的绝对数量来看，需要增加十倍甚至上百倍的链接才能与竞争对手相抗衡。限于亿赐客团队的规模，这是不现实的。为了更深入了解外部链接上的差距，我用 Majesty SEO 进行了查询。由于 Majesty SEO 不是免费工具，下面的抓图就不提是哪一个网站的数据了，只是为了让读者能够有一个大致的概念。

图 14-2 显示的是其中一个竞争对手网站外部链接按不同域名种类分类得到的数据。“Referring Domains”指的是总域名数，“External Backlinks”指的是总外部链接数。可以看到，虽然这些竞争对手外部链接总数都在十多万几十万以上，但其实链接来自于有限数目的域名，总域名数远远低于总外链数，通常是 2000～3000 个独特域名。相信其中一些域名网站规模大，给了全站链接，所以造成链接总数庞大。

#	Short TLD	Long TLD	Referring domains	External backlinks
1	au	.com.au	1	18
2	be	.be	1	2
3	bg	.bg	1	1
4	biz	.biz	6	19
5	ca	.ca	2	2
6	cc	.cc	33	598
7	ch	.ch	3	3
8	cn	.cn	766	42,711
9	cn	.com.cn	290	27,012
10	cn	.edu.cn	4	13
11	cn	.gov.cn	5	13
12	cn	.net.cn	25	71
13	cn	.org.cn	21	105
14	com	.com	2,072	836,514
15	de	.de	12	9,087
16	es	.es	1	2
17	eu	.eu	4	6
18	fi	.fi	1	3
19	fr	.fr	1	22,422
20	hk	.hk	5	13
21	hk	.com.hk	1	7
22	ie	.ie	1	1
23	in	.in	3	4
24	info	.info	32	57
25	io	.io	2	5
26	it	.it	6	7

图 14-2 某比较购物网站来自不同种类域名的外部链接

从 2000～3000 个不同域名吸引外部链接，看起来就变成一个有可能达到的目标。前面提到过，外部链接的来源域名数目是决定外链实力的主要依据之一，从某种程度上说，比外部链接绝对数量更重要。这个数据使我们增强了信心，外部链接建设方面还有机会拉近与竞争对手的距离。

14.1.3 Alexa 数据

Alexa 是站长都非常熟悉的流量排名服务，还可以查看到一些其他网站信息。图 14-3 是 5 个竞争对手的 Alexa 排名。

可以看到聪明点和丫丫网排名比较靠前，其他三个大致相同。不过种种迹象显示，丫丫网的流量应该没有达到聪明点的水平。Alexa 提供的数据并不准确，只能作为大致参考。

图 14-4 显示的是用户每次访问平均浏览的页面数。

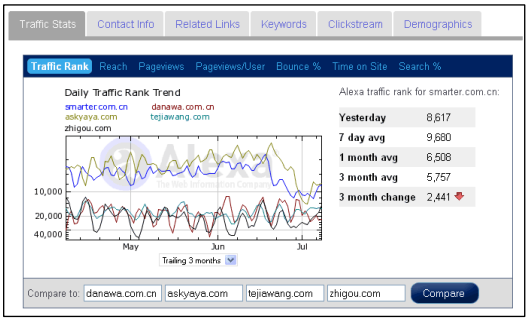


图 14-3　几个比较购物网站的 Alexa 排名

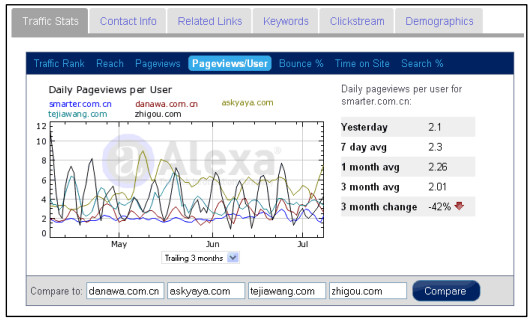


图 14-4　比较购物网站用户浏览页面数

智购网、丫丫网，浏览页面数起伏过大，没有什么参考意义。聪明点和大拿网的浏览页数都是 2 左右，这个数字相较其他类型网站显得略低，说明比较购物网站黏度都不高。不过这也符合比较购物网站的特点，其目的就在于把用户送到其他购物网站上。

图 14-5 显示的是跳出率。

同样，丫丫网的跳出率起伏过大，不具参考意义。其他几个网站的跳出率都在60%左右，这样的跳出率相较于正常购物网站就显得偏高。和用户浏览页面数一样，这大概也是比较购物网站的特性所决定的。

用户停留时间与浏览页面数有基本相同的趋势，用户平均在网站上停留不到 2 分钟，如图 14-6 所示。

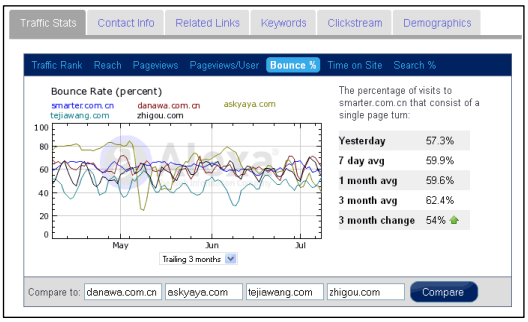


图 14-5　比较购物网站跳出率

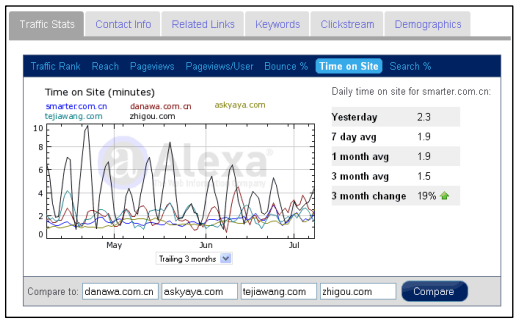


图 14-6　比较购物网站用户停留时间

图 14-7 显示的是搜索流量占所有流量的比例。

聪明点网站搜索流量比例最高，达到 40%～50%。从下面对网站页面的分析也可以看到，聪明点网站 SEO 做得很好。智购和大拿网搜索流量也占到 30%～40%，并且智购网的搜索流量比例在不断提高中。对网站的简单分析表明，智购网的 SEO 水平也很不错。

丫丫网搜索流量不到 10%，显得不合常理。加上页面浏览数和跳出率数据的大幅起伏，丫丫网流量可能包含一些很不常规的来源。

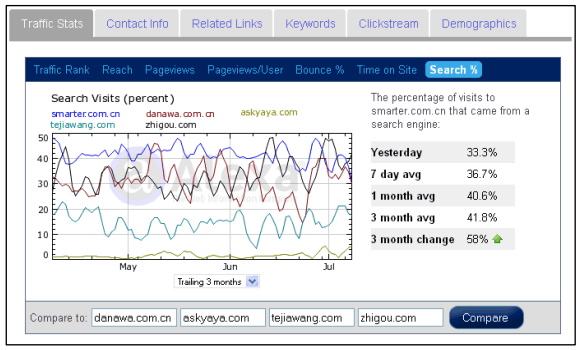


图 14-7　比较购物网站搜索流量

14.1.4　Google 趋势流量

Alexa 排名只显示网站的流量世界排名，并不显示具体流量数字，Google Trends 则能显示具体数字。图 14-8 是 2009 年 7 月份 Google Trends 所显示的 5 个竞争对手网站流量。

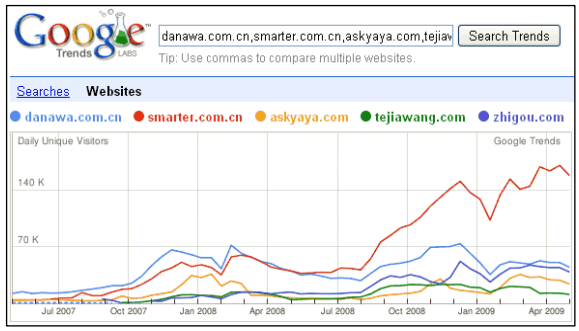


图 14-8　2009 年 7 月 Google Trends 显示的竞争对手网站流量

可以看到，聪明点网站流量最大，4 月份时达到日 IP 16 万左右。不过 2009 年 12 月份再用 Google Trends 查流量时看到，从 6 月份开始聪明点流量由于某种原因急剧下降，但还是保持在第一梯队中，如图 14-9 所示。

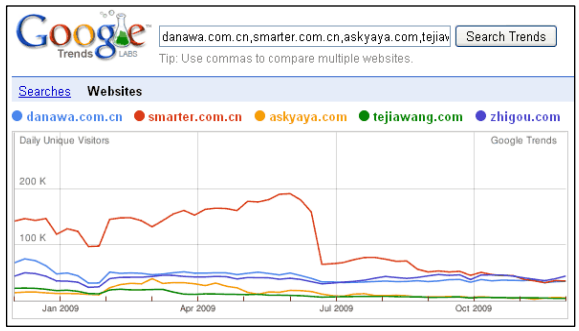


图 14-9　2009 年 12 月 Google Trends 显示的竞争对手网站流量

7 月份和 12 月份的流量数字都显示，聪明点、智购和大拿网流量应该在每天 5 万 IP 上下。特价王和丫丫网不到上面三个网站的一半。5 个竞争对手网站中，聪明点和智购没有使用 Google Analytics，丫丫、特价王、大拿网都使用了 Google Analytics，所以至少丫丫、特价王、大拿网的流量数字是相对准确的。通过相关行业人士侧面了解，当时第一梯队的比较购物网站真实流量与此数字相差不远。

14.1.5 网站品牌名称热度

用户搜索网站品牌名称的次数能很好地显示网站的知名度。使用 Google Trends 搜索亿赐客及主要竞争对手的品牌名称，情况如图 14-10 所示。

图 14-10　Google Trends 显示的名牌名称搜索趋势

搜索大拿网、聪明点和特价王的稍微多一点，但是曲线起伏过大，说明搜索次数绝对数值都比较小，一点点数量变化都会造成曲线的大幅波动。这个情况说明比较购物在中国还没有被普遍接受，更没有权威的、大家都知道的品牌。

Google Insights 服务也可以显示品牌名称搜索趋势，如图 14-11 所示。

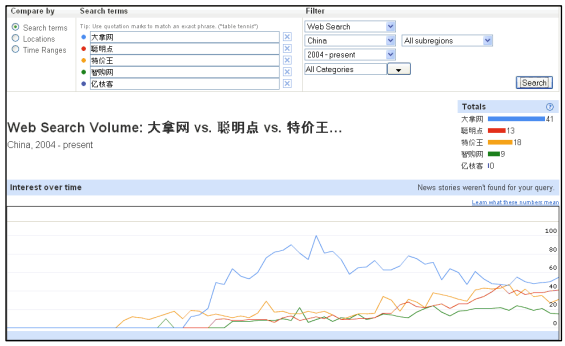


图 14-11　Google Insights 显示的名牌名称搜索趋势

曲线稍微平滑一些。可以看到大拿网搜索次数最多，但近期呈下降趋势。聪明点品牌关注度则稳步上升，是最值得关注的竞争对手。

以上两个图都显示亿赐客（当时名称为亿枝客）的品牌关注度近乎为零。

14.1.6 英文比较购物网站情况

我也简单查看了一下英文主要比较购物网站的情况，如图 14-12 所示。

图 14-12　2009 年 7 月 22 日英文主要比较购物网站 SEO 数据

同样，选出最重要的指标列成表 14-2。

表 14-2　2009 年 7 月 22 日记录的英文比较购物网站数据

	Google PR	首页快照新鲜度	年　　龄	美味书签	雅虎全站外链	雅虎首页外链	Google 收录	开放目录
PriceGrabber	7	无	1995 年 5 月	4712	8 320 000	1 260 000	2 170 000	收录
NexTag	7	2009 年 7 月 22 号	1999 年 10 月	1943	2 120 000	1 450 000	8 110 000	收录
BizRate	7	2009 年 7 月 20 号	1997 年 10 月	2238	1 350 0000	750 000	2 770 000	收录

再看这 3 个英文网站的流量，如图 14-13 所示。

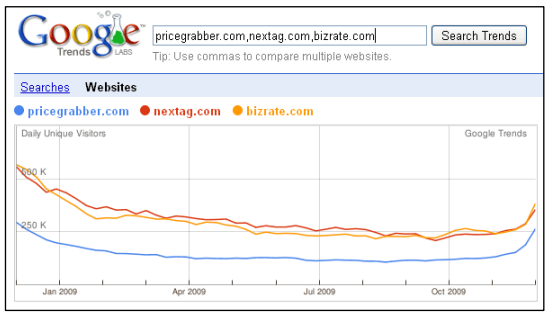


图 14-13　英文比较购物网站流量

这 3 个网站流量都相对稳定，只是在圣诞购物期间流量增长，没有像中文网站那么多的起伏。2009 年 NexTag 和 BizRate 网站的流量基本稳定在日 25 万 IP 以上。

从网站历史、收录、外部链接和流量看，英文购物网站的实力比中文购物网站要高

出一个数量级。尤其是外部链接，更是达到了上千万级别，而且链接分布非常自然，首页链接只占链接总数的很小一部分。

14.2 竞争对手网站研究

下面我们再来进一步观察竞争对手网站的更详细情况，尤其是页面及结构的优化。

从上面的大致数据可以看到，聪明点是比较强的竞争对手，所以单个网站的进一步分析就以聪明点为例。

14.2.1 域名注册信息

如图 14-14 所示，域名注册于 2002 年 11 月份，不算是一个很老的域名，但与亿赐客及其他同行业网站相比，已经算是比较有历史的域名了，占有一定先天优势。

图 14-14　聪明点域名注册信息

14.2.2 基本信息

还是使用 SEO for Firefox 插件。

图 14-15 是 2009 年 7 月 12 号的抓图，快照是 2 天前的。网站既有迷你全站链接，也有普通全站链接。

图 14-15　聪明点网站基本数据

14.2.3 外部链接

通过 Yahoo! Site Explorer 查看，全站外部链接为 100 万个以上，如图 14-16 所示。Yahoo! Site Explorer 链接工具大致是按链接的重要性排序的。排在前面的外链中可以看到不少权重相当高的网站，像阿里妈妈、美国聪明点总部网站等，如图 14-17 所示。

雅虎链接工具显示的首页链接有 92 万多个，说明绝大多数外部链接是指向首页的，与主流英文比较购物网站相比过于集中，分类和产品页面外部链接比较少。

图 14-16　聪明点网站外部链接

图 14-17　聪明点网站首页外部链接

14.2.4　收录

图 14-18 显示的是 2009 年 7 月聪明点网站在 Google 的收录数字，580 多万个页面，是一个有相当大规模的网站。百度收录数字则更高，达到 1000 万个以上，如图 14-19 所示。

图 14-18　聪明点网站在 Google 的收录

另外一个与收录有关的重要数据是网站近期收录的新页面数字。使用 Google 百宝箱可以看到，在过去 24 小时内，聪明点网站新增加收录页面 1290 条，如图 14-20 所示。

图 14-19　聪明点网站在百度的收录

图 14-20　过去 24 小时聪明点网站新收录页面

在过去一星期新收录 25 000 多个页面，如图 14-21 所示。

过去一年新收录 50 万个页面，如图 14-22 所示。

图 14-21　过去一星期聪明点网站新收录页面

图 14-22　过去一年聪明点网站新收录页面

网站新增加的收录页面数字能够在一定程度上体现出网站在搜索引擎眼中的权重。按照这个速度，聪明点网站每年增加 50 万个甚至更多的新页面，带来的长尾流量自然也会增加。

14.2.5　QQ 书签

SEO for Firefox 包含了美味书签、Digg 等常用英文社会化网站数据，这对中文网站来说并不太重要。为了查看竞争对手网站在社会化书签服务中的表现，我选择中文用户常用的 QQ 书签。

从图 14-23 可以看到，聪明点网站在 QQ 书签中被收藏过 300 次以上，其中大部分是首页，被收藏 296 次。

相比之下，大拿网被收藏 1000 多次，如图 14-24 所示。

图 14-23　聪明点在 QQ 书签中的表现

图 14-24　大拿网在 QQ 书签中的表现

丫丫网被收藏 60 多次。智购网被收藏 20 多次。亿赐客网站只有 1 次，如图 14-25 所示。

图 14-25　亿赐客在 QQ 书签中的表现

14.2.6　外链锚文字

启用 SEO Link Analysis 插件，然后在 Yahoo! Site Explorer 查外链，可以看到外部链接使用的锚文字，如图 14-26 所示。

可以看出聪明点网站并没有把注意力放在“比较购物”上，而是放在“网上购物”这个更通用、更热门的关键词上。外链锚文字是聪明点网站在搜索“网上购物”时排名不错的原因之一。图 14-27 是我在 2009 年 7 月 22 日时查的聪明点网站主要关键词排名。

2009 年 12 月再次查询时，“网上购物”这个词，聪明点网站在 Google 上升到第 12

位。外部链接的数量固然重要，使用的锚文字与排名也有直接关系。

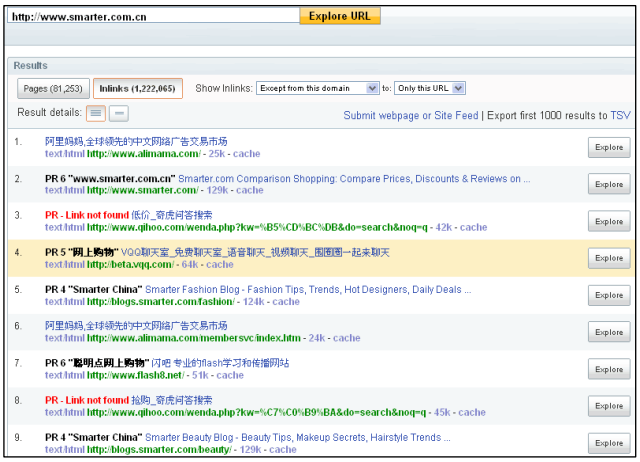


图 14-26　聪明点网站外部链接使用的锚文字

1	报表生成时间:2009-07-22 02:45:50		
2	查询网址:http://www.smarter.com.cn　查询范围:1-10页		
3			
4	关键字/搜索引擎	百度	Google
5	购物	在1-10页中未查询到结果	搜索排名为37
6	网上购物	搜索排名为53（推广4 - 自然49）	搜索排名为17
7	网购	在1-10页中未查询到结果	在1-10页中未查询到结果
8	比较购物	搜索排名为10（推广0 - 自然10）	搜索排名为2
9	导购网	在1-10页中未查询到结果	搜索排名为46
10	购物网站	在1-10页中未查询到结果	搜索排名为14
11	网上商城	在1-10页中未查询到结果	在1-10页中未查询到结果

图 14-27　聪明点网站主要关键词排名

做个比较，我们再看智购网站的外部链接锚文字，如图 14-28 所示。

可以看到智购网更是把目标放在“网上购物”这个关键词上，相应地，“网上购物”这个词在百度及 Google 排名都更靠前，如图 14-29 所示。

图 14-28　智购网站外部链接使用的锚文字

1	报表生成时间:2009-07-22 03:15:10		
2	查询网址:http://www.zhigou.com　查询范围:1-10页		
3			
4	关键字/搜索引擎	百度	Google
5	购物	在1-10页中未查询到结果	搜索排名为31
6	网上购物	搜索排名为15（推广4 - 自然11）	搜索排名为7
7	网购	在1-10页中未查询到结果	在1-10页中未查询到结果
8	比较购物	搜索排名为13（推广0 - 自然13）	搜索排名为9
9	导购网	在1-10页中未查询到结果	在1-10页中未查询到结果
10	购物网站	在1-10页中未查询到结果	在1-10页中未查询到结果
11	网上商城	在1-10页中未查询到结果	在1-10页中未查询到结果

图 14-29　智购网站主要关键词排名

考虑到智购网外部链接总数比聪明点网少得多，更可以看出锚文字的重要意义。

我们再看大拿网外部链接锚文字，如图 14-30 所示。

目标比较分散，其结果就是“网上购物”这类最热门关键词排名与智购和聪明点网有一定差距。

而丫丫网的外链锚文字则放在了“比较购物”上，如图 14-31 所示。

由于“比较购物”被搜索的次数不高，就算这个词排名在前面，带来的流量也十分有限。

图 14-30　大拿网外部链接使用的锚文字

图 14-31　丫丫网外部链接使用的锚文字

14.2.7　网站首页优化

判断一个网站 SEO 水平，除了看外部链接数据，当然同样重要的是看网站本身的页面及结构优化。有时候外部链接强大，只是因为公司是大品牌，不用 SEO 人员进行链接建设，公司的品牌、历史、公共关系就足以吸引巨量链接。从这个意义上来说，网站本身的优化更能体现出 SEO 水平。

下面我们就来简单看一下聪明点网站本身优化情况。

聪明点网站首页如图 14-32 所示。

总体上来说，看过首页源代码后，就可以知道这是一个经过专业 SEO 处理的网站，而且水平相当高。通过下面几点可以说明。

图 14-32　聪明点网站首页

1．Meta 标签

页面 Meta 标签部分代码是这样的：

```
<TITLE>网上购物/商品导购网/比价购物-聪明点比较购物网</TITLE>
<META name="description" content="聪明点比较购物网，致力于提供方便、详细地比较
相似商品的价格，规格参数，我们的产品信息达到 100 多万，是您网上购物时最佳的比价购物网，
我们的产品包括电脑硬件、数码产品、手机、汽车用品、办公用品、鲜花园艺、化妆品等。聪明
点帮您作出最聪明的网上购物选择。">
<META name="keywords" content="网上购物,比价购物,购物网,商品导购">
```

可以感觉到，页面标题是经过精心设计的，长度适当，最重要关键词“网上购物”

放在最前面，其他次要关键词，如比价购物、比较购物、商品导购，各出现一次。整个标题突出了关键词而不堆积。

描述标签虽然对排名没什么影响，但对吸引点击有不小作用。聪明点首页的说明标签写得也很自然，既包含了相应关键词，又突出了方便、详细、数据量大等特点，吸引用户点击。

2．H1 标签

H1 标签用在页面右上角公告栏，文字“聪明点比较购物更省钱更省心”的地方，锚文字只包含“聪明点比较购物”。相应的 HTML 代码是：

```
<DIV class="content_gonggao">
   <H1>聪明点比较购物</H1>，更省钱，更省心！<BR>
   千家网上商店，近 300 万在线商品！
</DIV>
```

右上角公告栏处的这一行文字从视觉上看与普通文字没有明显差异，很难看出会是标题标签，这显然是 SEO 的结果。

聪明点网站的 H1 并没有放上最重要的关键词“网上购物”。其原因我不知道，但并不能说明首页没有优化好。外人看一些大网站，往往并不能知道很多细节处理的真正原因。前面提到，本节的详细 SEO 案例我不会去分析其他网站，哪怕是 SEO 做得非常好的网站，因为外人不知道呈现出来的代码具体原因是什么，就像这个地方。

通常的观点是一定要把“网上购物”这个词放进 H1。但聪明点 SEO 人员没有放，有各种可能的原因。比如：

- 他们不想使页面过度优化。
- 因为用户体验，文字“聪明点比较购物更省钱更省心”，比“聪明点网上购物更省钱更省心”，更能说明网站的功能。
- 在做测试，也可能就是测试的结果。
- 网站流量依靠长尾，热门关键词带来的流量有限，不必太优化。
- SEO 团队认为 H1 标签作用很小。

总之，判断一个网站的 SEO 情况，不能从单一细节出发，而要看整体。某个特定细节很可能显得不够优化，甚至是有问题的，但外人并不了解之所以这样的真正原因。

3．H2 使用

页面上也适当使用了 H2 标签，比如左侧导航的分类文字就是放在 H2 中的，HTML 代码如下：

```
<!--导航开始-->
            <DIV class="sidebox">
            <DIV id="nav">
```

```
                <UL>
                  <LI class="line"><A
href="http://www.smarter.com.cn/ cosmetics-693/"><H2>化妆品
</H2></A></LI>
                  <LI class="line"><A
href="http://www.smarter.com.cn/ shoessuitcasesbags-1018/"><H2>鞋子箱包
</H2></A></LI>
                  <LI class="line"><A
href="http://www.smarter.com.cn/baby-1014/"><H2>母婴用品</H2></A></LI>
```

页面中间的小标题，如“热卖商品”，也是放在 H2 中，这与整个页面的语义结构相符合。HTML 代码如下：

```
<DIV class="title">
   <H2>热卖商品</H2>
</DIV>
```

4．文案写作

页面导航和产品列表、新闻信息等都是来自于数据库，而且大部分应该是经常自动更新的，其文字比较难以从 SEO 角度进行优化。聪明点首页特意拿出一块只适用于首页的固定的地方，放上与主要关键词高度相关的文字，这就是左侧导航下面的“聪明点宣言”，如图 14-33 所示。

通过这段文字，文案写作人员简明扼要地告诉用户为什么应该使用聪明点网站的服务。同时，在用户体验的基础上自然融入相关词汇，适当地提到了相关关键词，如商品、商家、产品、网购、在线等，这些词都是与“网上购物”语义相关的，对网上购物这个词起到支持作用。

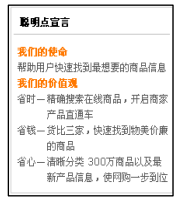


图 14-33　聪明点网站首页宣言

5．URL 静态化

整个网站的 URL 完全静态化，包括分类页面、产品页面、产品列表页面上的翻页等，几乎所有页面都做了静态化处理。更难得的是连搜索页面也做了静态化，如页面顶部“热门搜索”中，每个搜索词对应的链接都是静态化 URL，其 HTML 代码为：

```
<STRONG>热门搜索：</STRONG><A
href="http://www.smarter.com.cn/_nemllgkpnakm -se-ch-1008-c-6/">运动鞋
</A> <A
href="http://www.smarter.com.cn/sharp-se -ch-666-c-667/">Sharp 手机
</A> <A
href="http://www.smarter.com.cn/ _miljnhnd-se-ch-1016-c-6/">裙子
</A> <A
href="http://www.smarter.com.cn/ communication-666/category/threeg-671/
">3G 手机</A> <A
href="http://www. smarter.com.cn/_lhmamjlj-se-ch-693-c-695/">防晒
</A> <A
href="http://www. smarter.com.cn/_mglllljpl-se-ch-642-c-647/">苹果播放器
</A> <A
href="http: //www.smarter.com.cn/_nfkfngknllpk-se-ch-1015-c-19/">榨汁机
```

```
</A> <A href= "http://www.smarter.com.cn/auto-1020/prod-83876/">一
汽马自达</A>    <A
href="http://www.smarter.com.cn/top-searches
.htm">更多...</A>
```

像这么完整 URL 静态化的大型网站并不多见。

6．NoFollow 的使用

整个页面上多次使用 NoFollow 属性，并且使用目的明确。比如热卖商品部分的产品链接，如图 14-34 所示。

图 14-34　聪明点首页产品链接

相应 HTML 代码如下：

```
<LI class="border_bottom ">
           <SPAN class="pro_boximg"><A href="
http://www.smarter.com.cn/ gift-1010/prod-704564/"
target="_blank"><IMG src="./704564.jpg" alt="瑞士军刀 Victorinox 维氏-瑞士
卡（..."></A></SPAN>
                <STRONG><A href="http://www.smarter.com.cn/gift-1010/
prod-704564/" target="_blank">瑞士军刀 Victorinox 维氏-瑞士卡（...</A></STRONG>
                <SPAN class="blue">
                  <A href="http://www.smarter.com.cn/gift-1010/prod-
704564/" target="_blank" rel="nofollow">
                                          ￥125.00 - ￥179.00
                                    </A>
                </SPAN>
                <DIV class="but">
                     <A href="http://www.smarter.com.cn/gift-1010/prod-
704564/" target="_blank" rel="nofollow">
                                                  <IMG src="./but_price.gif"
alt="比较价格">
                                        </A>
               </DIV>
</LI>
```

可以看到，产品图片、名称、价格和“比较价格”按钮都是指向产品页面的链接，但其中价格和比较价格按钮使用了 NF 属性，使搜索引擎蜘蛛不要跟踪这两个链接，其意义就在于这两个链接的锚文字对产品页面没有任何帮助。第一个链接中的图片 ALT 文字和第二个链接中的锚文字都可以使对应的产品页面相关性提高，改善排名。而“￥125.00 - ￥179.00”这样的文字，和比较价格按钮中的 ALT 文字“比较价格”，对产品页面没有任何说明意义。这两个链接如果允许搜索引擎跟踪，并把锚文字计入算法中，只能冲淡相关性。

顺便提一下，在查看聪明点首页时，曾有一个迷惑不解的地方，就是热门产品下的 6 个产品，有 4 个使用“比较价格”作为按钮上的文字，有 2 个使用“产品详细”作为按钮上的文字（请参考图 14-32），不知道意义在哪里，可能是某种测试看哪个点击率比较高。

后来多次查看首页发现，不同按钮文字的数目、位置并不固定，图 14-35 就显示了另一种布局。后经读者指点才明白，“产品详细”对应页面是只有一个商家在卖的产品，不存在“比较价格”问题。

图 14-35　聪明点首页产品按钮

左侧的商家排行链接中使用了 NF，代码是这样的：

```
<DIV class="buss_top">商家降价幅度排行</DIV>
   <UL>
                        <LI class="lileft">1</LI>
                        <LI class="limid"><A
href="http://www.smarter. com.cn/goldenbook-5954/mcate-1011-0/">中国科技
金书网</A></LI>
                        <LI class="liright"><STRONG><A
href="http://www. smarter.com.cn/goldenbook-5954/mcate-1011-0/"
rel="nofollow">179731</A></STRONG></LI>
                     </UL>
                                                <UL>
```

同样，第一个链接使用商家本身的名称为锚文字，增加了相关性。第二个链接文字是一个数字，应该是这个商家的产品总数，对商家页面没有任何语义上的相关性，所以用 NF 属性阻止搜索引擎跟踪和计算锚文字。

页面底部链接也有一些使用了 NF 属性，比如“免责声明”、“联系我们”。显然，这些页面在搜索意义上并不重要，通常就不应该分散权重到这些页面。而像“网站地图”、“友情链接”这些页面对 SEO 有很大意义，通常不使用 NF。

14.2.8　其他页面优化

我们再来简单看一下其他页面。

1．一级分类页面

如图 14-36 所示为化妆品分类页面。

面包屑导航、左侧分类都很清晰，有利于用户浏览和搜索引擎抓取。右上角的推荐品牌列表和左侧导航部分的推荐商家，都链接到相应的品牌及商家产品页面，这是常用的增加页面收录入口的手法。

图 14-36　聪明点一级分类页面

有意思的 SEO 技巧出现在页面底部，“化妆品简介”部分，如图 14-37 所示。

图 14-37　聪明点一级分类页面底部说明文字

通常电子商务网站的分类页面比较不容易提供独特的原创内容。大多数电子商务网站都是从数据库中提取热门产品列在分类页面上，这造成不同网站之间及网站本身不同页面之间内容相似。为了缓解这个问题，聪明点网站给所有分类页面，包括一级分类和下面要看的二级分类，都写了一段文字说明。这段说明文字在有的页面上是使用模板插入分类名称自动生成的，有的是根据不同分类的特点写的独特文字，如图 14-37 所示的“香水简介”。

2．二级分类页面

产品列表翻页，不同的排列方式（如受欢迎程度、商品评分、价格排序），不同的显示方式（列表显示、栅格显示），所有 URL 全部静态化，如图 14-38 所示。

比较有意思的 SEO 手法也是出现在底部，如图 14-39 所示。

图 14-38　聪明点二级分类页面

图 14-39　聪明点二级分类页面底部文字及链接

除了“香水简介”部分的原创文字之外，页面还列出了热门搜索和热门类别，这些链接为相应的页面提供了更多收录入口，提高整个网站的收录比例。

3．产品页面

产品页面上的安排与大多数电子商务网站类似。在这里我发现了一个有意思的现象，几个主要竞争对手网站都使用了这个手法，那就是把产品分成不同的页面，也就是用主产品页面显示产品说明及价格比较，如图 14-40 所示。

用户点击“产品参数”或“发表评论”这两个选项卡后，将被带到不同的页面。如图 14-40 所示，产品主页面 URL 是：

http://www.smarter.com.cn/cosmetics-693/prod-70740478/

而产品参数页面 URL 是：

http://www.smarter.com.cn/cosmetics-693/pspec-70740478/

如图 14-41 所示。

图 14-40　聪明点主产品页面

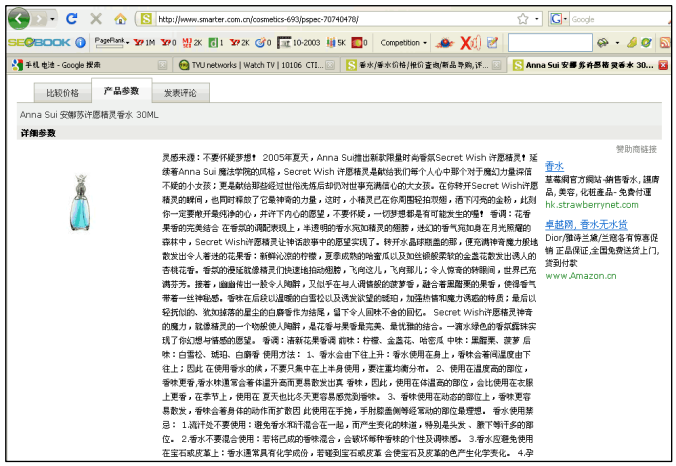


图 14-41　聪明点产品参数页面

发表评论页面 URL 是：

http://www.smarter.com.cn/cosmetics-693/pwrite-70740478/

如图 14-42 所示。

通常使用选项卡的目的是为了节省页面空间，用户点击选项卡后显示相应内容，但

页面并不改变 URL，也不重新调入。聪明点和其他几个中文比较购物网站把产品参数和发表评论这样的选项卡安排成不同的页面、不同 URL，我猜想其目的在于增加网站总页面数，使网站看起来规模更大，也就使权重增加。另外，不同页面也方便针对不同关键词，比如主页面关键词是“产品名称+价格/报价”，产品参数页面的主打关键词就变成了“产品名称+参数”，发表评论页面主打关键词变成了“产品名称+评论”。

图 14-42　聪明点产品评论页面

在 Google 搜索一下就会看到这几个页面确实都有收录，如图 14-43 所示。

图 14-43　聪明点产品页面收录

应该说这是一个不错的点子，不过并不适用于所有网站。在亿赐客这个具体案例中，我倾向于不要学习这种方式。

综合外部链接及网站本身优化的情况看，聪明点是一个经过专业 SEO 的网站，而且水平相当高。其他几个竞争对手网站优化得也都不错，尤其是智购，与聪明点不相上下。

14.3　亿赐客网站分析

本节看一下亿赐客网站在 2009 年 7 月优化前的情况。

14.3.1 域名注册

如图 14-44 所示，域名注册时间为 2007 年 8 月。虽然算不上一个全新的网站，但与其他竞争对手网站，或广泛点说，与其他电子商务网站相比，算是年轻的。所以域名权重较难在短期内有质的飞跃，除非能自然获得大量高质量外链，这种外链可遇而不可求，外链建设任重道远。

Domain names in the .com and .net domains can now be registered with many different competing registrars. Go to http://www.internic.net for detailed information.

Domain Name: YICIKE.COM
Registrar: XIN NET TECHNOLOGY CORPORATION
Whois Server: whois.paycenter.com.cn
Referral URL: http://www.xinnet.com
Name Server: NS.XINNET.CN
Name Server: NS.XINNETDNS.COM
Status: ok
Updated Date: 17-nov-2008
Creation Date: 21-aug-2007
Expiration Date: 21-aug-2010

图 14-44　亿赐客域名 whois 信息

14.3.2 Google PR 值

在 2009 年 6 月 23 日 Google 工具条 PR 更新中，刚刚从 4 升到 5。PR 值并不是线性的，PR 值 1 和 2 之间的差距，比 PR 值 4 和 5 之间的差距要小得多。PR 值越高，要想提升就越困难。亿赐客网站应该还处于 PR 值 5 的低端，与竞争对手的普遍 PR 值 6 相比，有一些差距。

14.3.3 收录

图 14-45 所示是在 Google 的收录情况。

图 14-45　亿赐客网站在 Google 的收录

根据技术人员的统计，2009 年 7 月份网站已经有上百万的产品数据，而 Google 只收录了 7 万多个页面，说明收录非常不充分。通常收录不好的原因无外乎几个：网站权重太低，网站内容原创度太低，网站架构不利于收录。由于网站 PR 值和权重及内容本身不容易迅速改进，改善网站架构就成了提高收录的主要方法。

单纯 site:指令的准确性较低，我也经常使用另一个方法估算网站被收录页面数，就是搜索一个在其他网站上不会有、而在你的网站上每个页面都会有的字符串，比如公司电话、地址、电子邮件地址、口号等。对中文网站来说，一个很好的独特字符串是正常情况下（如不存在多个网站共用一个备案号）的备案号码。搜索亿赐客网站备案号，如

图 14-46 所示，Google 显示的是 45 000 多个页面，与 site:指令得到的数字在同一量级，说明亿赐客网站 7 月份的收录数大致不会超过 10 万。也就是说，收录率在 10%以下，有很大的改进空间。图 14-47 是在百度的收录数据抓图。

图 14-46　在 Google 搜索亿赐客备案号

图 14-47　亿赐客网站在百度的收录

在百度收录 335 000 个页面，比在 Google 中表现稍微好一点，但是 30%多的收录率还属于比较低的。

我们再来看最新收录的页面数据。使用 Google 百宝箱按时间显示功能，可以看到亿赐客网站在最近 24 小时收录 550 个页面，如图 14-48 所示。

过去一个星期收录 4170 个新页面，如图 14-49 所示。

图 14-48　过去 24 小时收录新页面

图 14-49　过去一星期收录新页面

在过去一年收录了 2 万多个新页面，如图 14-50 所示。

与前面看过的聪明点网站相比，这个收录速度就相差太远了。如果在收录上不能有质的改变，收录页面差距将越拉越大，而收录直接决定了长尾流量的多少，是这种中型网站优化最最重要的一个方面，必须解决。

图 14-50　过去一年收录新页面

14.3.4　外部链接

如图 14-51 所示，在 Yahoo! Site Explorer 查询，网站全站链接 2 万多一点，排在前面的没有什么著名网站，至少没有我所知道的有名网站。

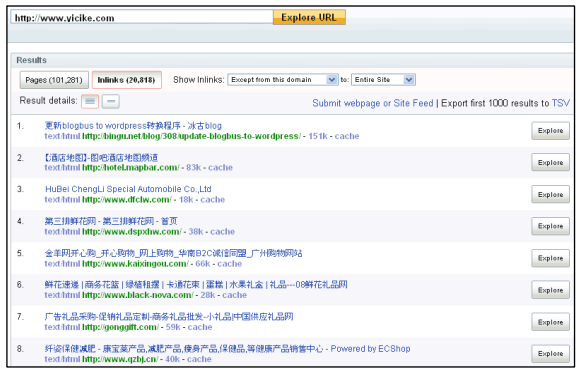


图 14-51　亿赐客网站全站外链

首页的链接 18 000 多个，也就是说 90%以上的外部链接是集中在首页，整个网站的外部链接构成不是很理想，缺少指向内页的链接，如图 14-52 所示。

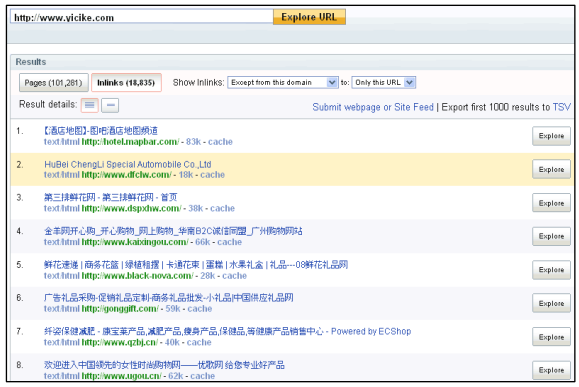


图 14-52　亿赐客网站首页外链

我们再看外部链接锚文字，如图 14-53 所示。

图 14-53　亿赐客网站外链锚文字

显然亿赐客一直以来是把“比较导购”作为首页的目标关键词。比较导购虽然是这类网站的相对准确功能描述，但是搜索数量太少，从 SEO 角度来看基本没有什么价值，不适合作为主要目标关键词。以后在交换链接或其他链接建设时，都需要把目标锚文字改到重新定位的主关键词上。

14.3.5　QQ 书签

如图 14-25 所示，网站在 QQ 书签只有 2 个收藏，其中一个还是后台管理，应该是亿赐客公司内部人员的书签。

14.3.6　基本流量数据

现在我们看一下 Google Analytics 显示的网站基本流量数据。

网站流量分析是一门学问，很深的学问。分析流量，找出问题，改进网站，往往是投入产出比最高的网络营销活动。

SEO 只是能从流量分析中获益的一小部分，流量分析的作用远远超出 SEO 范围，对于改进用户体验、提高转化率、提高销售都有重大意义。这些超出 SEO 范畴的内容在这里就不深入讨论了。下面只是向大家展示亿赐客网站优化前的基本流量数据，让读者有个初步认识，以便作为比较的基准。

1．流量概况

图 14-54 是 Google Analytics 控制台抓图。

图 14-54　Google Analytics 控制台及流量概况

平均日 IP 为 2000 多。对一个有百万数据的网站来说，这个流量表现不能令人满意，应该有很大的增长空间。

跳出率 83.48%，平均网站停留时间 1 分 1 秒，用户每次访问 1.55 个页面，这三个数据都表示网站黏度不高，但是与同行业网站比较也不算太差，黏度比较低是比较购物网站的共同特征。

2．流量来源

图 14-55 显示网站流量来源。

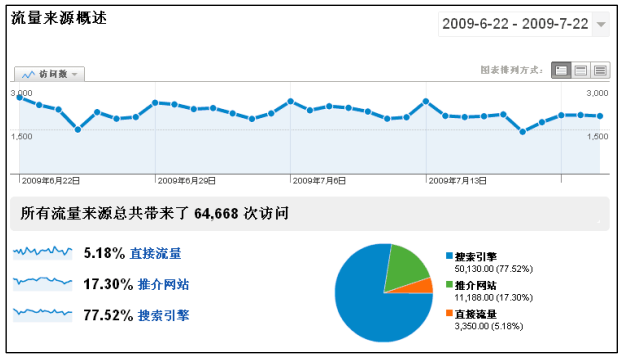


图 14-55　流量来源分布

搜索流量占所有流量的 77.52%，如果统计准确属正常范围，稍微偏高。真实搜索流量的比例应该比这个数字还要高，从后面的数据分析可以看到。相对来说，网站对搜索引擎的依赖已经比较大，这是潜在风险之一。如果可能，应该开拓用户重复访问及来自其他网站的点击流量。

搜索流量细节如图 14-56 所示。

来自百度和 Google 的流量大致相当。由于百度在国内市场占有率远高于 Google，

这样的搜索流量比例说明在百度的排名表现有更大问题，也就有更多的改善空间。从用户访问页数和停留时间看，Google 流量黏度稍微高于百度流量，但差距不是很大。

图 14-56　搜索流量细节

来自雅虎和必应的搜索流量黏度比百度、Google 都要高出不少，很可惜这两个搜索引擎的市场份额太低，可以忽略不计。

我们再看一下所有流量来源黏度，如图 14-57 所示。

图 14-57　所有流量来源黏度

其中标为 organic 的是自然搜索流量，标为 referral 的是来自其他网站的点击流量，标为 direct 的是直接访问（如用户从浏览器书签访问网站）。可以看到，直接访问流量黏度通常高于其他流量来源，这很符合逻辑，绝大部分网站都如此。

图 14-58 显示的是所有流量来源的目标转化率。

网站使用 目标转化

访问数	目标 1：搜索	目标 2：注册	目标 3：购物按钮	目标 4：广告	目标转化率	每次访问目标价值
64,668 网站总额百分比：100.00%	2.56% 网站平均数：2.56% (0.00%)	0.00% 网站平均数：0.00% (0.00%)	0.00% 网站平均数：0.00% (0.00%)	0.00% 网站平均数：0.00% (0.00%)	2.56% 网站平均数：2.56% (0.00%)	¥0.00 网站平均数：¥0.00 (0.00%)

	来源/媒介	访问数	搜索	注册	购物按钮	广告	目标转化率	每次访问目标价值
1.	baidu / organic	26,287	1.43%	0.00%	0.00%	0.00%	1.43%	¥0.00
2.	google / organic	23,596	2.52%	0.00%	0.00%	0.00%	2.52%	¥0.00
3.	sogou.com / referral	3,522	1.11%	0.00%	0.00%	0.00%	1.11%	¥0.00
4.	soso.com / referral	3,464	3.20%	0.00%	0.00%	0.00%	3.20%	¥0.00
5.	(direct) / (none)	3,350	9.91%	0.00%	0.00%	0.00%	9.91%	¥0.00
6.	baidu.com / referral	2,493	1.76%	0.00%	0.00%	0.00%	1.76%	¥0.00
7.	image.baidu.com / referral	299	1.67%	0.00%	0.00%	0.00%	1.67%	¥0.00
8.	bbs.yicike.com / referral	180	17.22%	0.00%	0.00%	0.00%	17.22%	¥0.00
9.	yahoo / organic	140	10.71%	0.00%	0.00%	0.00%	10.71%	¥0.00
10.	youdao.com / referral	113	4.42%	0.00%	0.00%	0.00%	4.42%	¥0.00
11.	bing / organic	87	13.79%	0.00%	0.00%	0.00%	13.79%	¥0.00
12.	web.gougou.com / referral	65	1.54%	0.00%	0.00%	0.00%	1.54%	¥0.00
13.	localhost / referral	61	11.48%	0.00%	0.00%	0.00%	11.48%	¥0.00
14.	cache.baidu.com / referral	59	6.78%	0.00%	0.00%	0.00%	6.78%	¥0.00
15.	aizhigu.com / referral	53	1.89%	0.00%	0.00%	0.00%	1.89%	¥0.00
16.	114search.118114.cn / referral	50	64.00%	0.00%	0.00%	0.00%	64.00%	¥0.00

图 14-58　所有流量的目标转化率

亿赐客网站把用户在网站上完成一次搜索设定为一次转化。在后面的优化报告中我会提到，这并不是一个好的转化目标，不过也在一定程度上说明了用户在网站上的互动和参与度，也有一定意义。Google 流量参与度比百度流量稍微高一点，雅虎和必应的流量质量高出更多。

排在第三位和第四位的分别是搜狗和搜搜，Google Analytics 把它们列为点击流量（referral），但其实这两个流量中的绝大部分也都应该是搜索流量。

图 14-59 显示的是搜狗流量来源细分。

网站使用 目标转化

访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
3,522 网站总额百分比：5.45%	1.30 网站平均数：1.55 (-16.22%)	00:00:35 网站平均数：00:01:01 (-41.80%)	97.33% 网站平均数：93.02% (4.63%)	83.30% 网站平均数：83.48% (-0.21%)

	维度：推介路径	访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
1.	/sogou	2,153	1.31	00:00:35	97.49%	83.42%
2.	/web	787	1.28	00:00:33	96.70%	83.10%
3.	/sohu	545	1.29	00:00:38	97.98%	82.94%
4.	/websearch/corp/search.jsp	27	1.15	00:00:34	88.89%	88.89%
5.	/express/sq.jsp	7	1.43	00:00:10	100.00%	85.71%
6.	/websnapshot	3	1.33	00:00:00	100.00%	66.67%

图 14-59　搜狗流量来源细分

从来源 URL（http://www.sogou.com/sogou?是搜索结果页面 URL 的最前一部分）可以看出，其实就是搜狗网站上的搜索流量。搜搜流量也是如此。这也就是为什么前面提到，搜索流量的比例实际高于 Google Analytics 显示的数字。Google Analytics 毕竟主要服务于英文网站，可能没有深入了解所有中文搜索引擎，没有把搜狗、搜搜这些小一点的中文搜索引擎流量算为搜索流量，而是算成了点击流量。

3．用户忠诚度

下面几个抓图显示的都是用户忠诚度或黏度。

首先是用户访问次数，如图 14-60 所示。

导出 电子邮件 添加到控制台 高级群体：所有访问次数

概述 »

访问者忠诚度 2009-6-22 - 2009-7-22

大多数访问的重复次数：1 次

来自此访问者的访问次数（包括当次访问）	访问者访问了 N 次的访问的次数	占所有访问的百分比
1 次		
20090622-20090722	60,195.00	93.08%
2 次		
20090622-20090722	2,093.00	3.24%
3 次		
20090622-20090722	401.00	0.62%
4 次		
20090622-20090722	184.00	0.28%
5 次		
20090622-20090722	114.00	0.18%
6 次		
20090622-20090722	70.00	0.11%

图 14-60　用户访问次数

93%的用户只访问网站一次，回头客太少。如果可能，增加互动性质的内容或者电子杂志，吸引用户多次访问，降低风险。

用户访问持续时间如图 14-61 所示。

访问时长 2009-6-22 - 2009-7-22

大多数访问持续的时间：0-10 秒

访问持续时间	这一时段的访问次数	占所有访问的百分比
0-10 秒		
20090622-20090722	54,754.00	84.67%
11-30 秒		
20090622-20090722	1,981.00	3.06%
31-60 秒		
20090622-20090722	1,770.00	2.74%
61-180 秒		
20090622-20090722	2,506.00	3.88%
181-600 秒		
20090622-20090722	1,981.00	3.06%
601-1,800 秒		
20090622-20090722	1,234.00	1.91%
1,801+ 秒		
20090622-20090722	442.00	0.68%

图 14-61　用户访问持续时间

84.67%的人只访问了 10 秒以内就离开网站，如果是信息类或普通电子商务网站，这是个过短的时间，用户没看什么内容就离开了，说明网站可能在视觉设计或易用性方面有问题。由于比较购物网站有一定的特殊性，与搜索引擎有点类似，这种网站的目标就是把用户送到其他网站，所以网站访问时间短并不一定是件坏事，还要看网站转化率如何。

页面跳出率如图 14-57 所示，网站整体跳出率为 83%，不太理想。其原因主要在于最终产品页面弹出率比较高，其实首页和分类页面弹出率还算不错。比如图 14-62 显示的是图书分类页面用户进入及离开的情况。

我们可以看到，90.44%的用户其实是点击访问了后续页面，表现还算不错。所以从网络营销角度看，真正要解决的是产品页面跳出率。

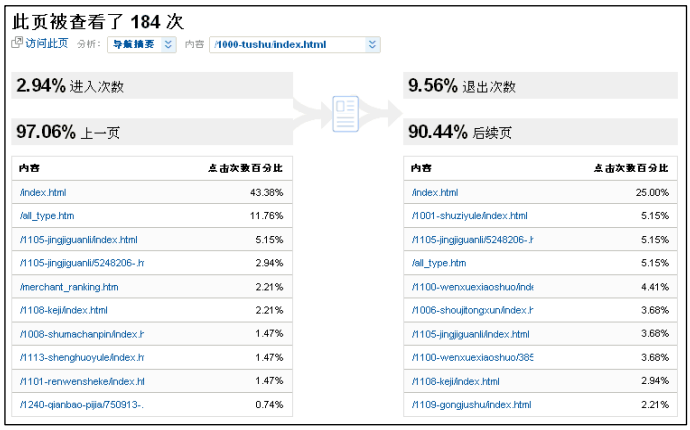


图 14-62　分类页面跳出率

和上面说的用户忠诚度、黏度一样，跳出率对比较购物网站来说也有它的特殊考虑，必须综合转化率来看。

4．用户地理位置

图 14-63 显示了用户地理位置。

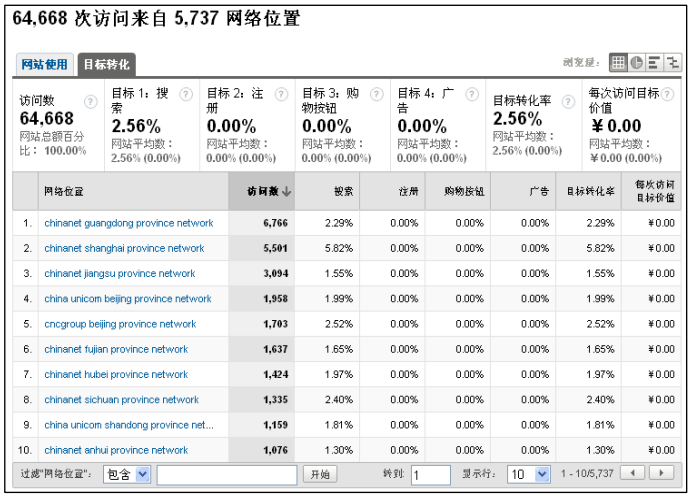


图 14-63　用户地理位置

其中一个数字比较奇怪。来自上海的用户目标转化率明显高于其他任何地方，黏度数字也同样如此。由于网站服务器位于上海，我曾经以为是因为服务器速度问题导致其他省市用户访问起来太慢，影响用户体验，所以转化率、黏度大幅下降，但后来通过流量细分发现不是这个原因。

细分是流量分析的重要方法之一，像前面提到的按不同搜索引擎查看流量质量、转化就是细分的一种。

Google Analytics 提供了一些简单细分工具，Google Analytics 把它称为群体。

如图 14-64 所示，站长可以查看 Google Analytics 已经做好的细分群体流量，也可以自定义其他细分群体。

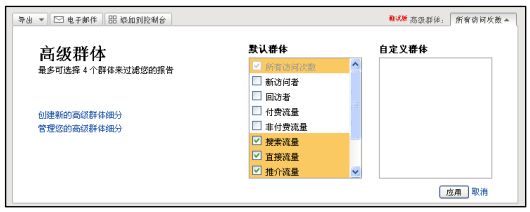
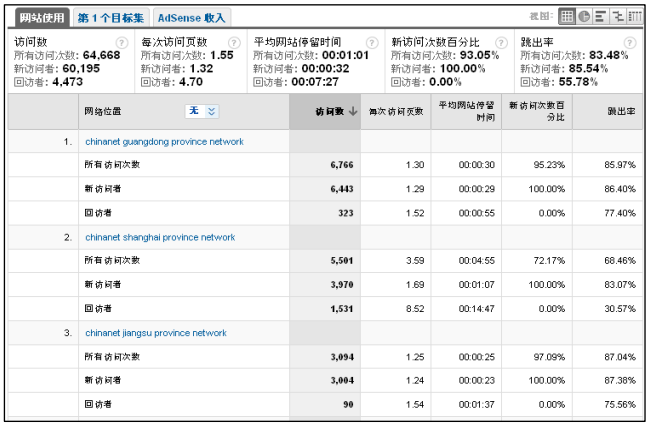


图 14-64　使用 Google Analytics 提供的群体进行流量细分

我在查看新访问者和回访者（也就是老用户）的地理位置数据时，发现上海地区用户新访问者与其他地区新访问者没有那么大的区别，但回访者的黏度远远高于其他省市回访者的黏度，如图 14-65 所示。



网站使用　第 1 个目标集　AdSense 收入

访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
所有访问次数: 64,668	所有访问次数: 1.55	所有访问次数: 00:01:01	所有访问次数: 93.05%	所有访问次数: 83.48%
新访问者: 60,195	新访问者: 1.32	新访问者: 00:00:32	新访问者: 100.00%	新访问者: 85.54%
回访者: 4,473	回访者: 4.70	回访者: 00:07:27	回访者: 0.00%	回访者: 55.78%

	网络位置	访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
1.	chinanet guangdong province network					
	所有访问次数	6,766	1.30	00:00:30	95.23%	85.97%
	新访问者	6,443	1.29	00:00:29	100.00%	86.40%
	回访者	323	1.52	00:00:55	0.00%	77.40%
2.	chinanet shanghai province network					
	所有访问次数	5,501	3.59	00:04:55	72.17%	68.46%
	新访问者	3,970	1.69	00:01:07	100.00%	83.07%
	回访者	1,531	8.52	00:14:47	0.00%	30.57%
3.	chinanet jiangsu province network					
	所有访问次数	3,094	1.25	00:00:25	97.09%	87.04%
	新访问者	3,004	1.24	00:00:23	100.00%	87.38%
	回访者	90	1.54	00:01:37	0.00%	75.56%

图 14-65　用户地理位置数据再按新旧用户细分

这说明问题在于上海地区回访者，而不在于上海地区的所有用户。考虑到亿赐客公司就在上海，基本可以断定这部分上海地区的回访者绝大部分是公司员工，所以黏度才极高，并拉高了上海用户的整体数字。这从上海地区回访者比例远远高出其他省市回访者比例也可以得到验证。

5．带来流量的关键词

图 14-66 显示的是百度带来流量的前 15 个关键词。

图 14-67 显示的是百度流量中靠后一些的关键词。

baidu 通过23,278 关键字 发出了 26,287 次 合计 访问

显示： 合计 | 付费 | 非付费

网站使用 目标转化 浏览量：

访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
26,287 网站总额百分比：40.65%	1.31 网站平均数：1.55 (-15.56%)	00:00:35 网站平均数：00:01:01 (-42.03%)	95.08% 网站平均数：93.02% (2.21%)	87.09% 网站平均数：83.48% (4.32%)

	维度： 关键字	访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
1.	site:www.yicike.com	147	12.31	00:22:40	2.04%	21.77%
2.	亿技客	126	3.29	00:04:33	63.49%	37.30%
2.	爱仕达厨具六件套刀具	103	1.32	00:00:17	97.09%	78.64%
4.	www.yicike.com	57	3.70	00:04:57	5.26%	45.61%
5.	多普达p2222	50	1.60	00:00:54	98.00%	66.00%
6.	礼想家多功能厨宝	43	1.07	00:01:07	97.67%	95.35%
7.	motoe6报价	33	1.21	00:00:11	100.00%	87.88%
8.	索尼 sony srs-a212 (白色 内置口径66mm喇叭单...	24	1.12	00:00:11	4.17%	87.50%
9.	motoem330	23	1.39	00:01:25	100.00%	73.91%
10.	motok1报价	23	1.22	00:00:06	91.30%	86.96%
11.	诺基亚6300c报价	21	1.00	00:00:00	100.00%	100.00%
12.	荷福磨力原图	20	1.25	00:00:37	75.00%	85.00%
13.	富士8100fd报价	18	1.17	00:00:12	88.89%	88.89%
14.	318220-102	17	1.12	00:00:09	88.24%	88.24%
15.	苹果翻盖手机报价	16	1.00	00:00:00	100.00%	100.00%

图 14-66　百度搜索流量的前 15 个关键词

32.	光泉冲凉踏步机	11	1.91	00:01:08	100.00%
33.	哈尔滨冰纯啤酒价格	11	1.18	00:00:08	100.00%
34.	森森水族箱报价	11	1.18	00:00:10	100.00%
35.	蜀虎一洗黑洗染香波	11	1.73	00:00:47	100.00%
36.	马连奴铁包价格	11	1.09	00:00:14	100.00%
37.	324777-007	10	1.20	00:02:13	70.00%
38.	great wall 长城 m172 17英寸 液晶显示器	10	10.30	00:28:06	0.00%
39.	vzi氧疗面膜素价格	10	1.00	00:00:00	100.00%
40.	www.flyco.com	10	1.10	00:00:03	100.00%
41.	解通价格	10	1.00	00:00:00	100.00%
42.	西施兰豆蔻价格	10	1.10	00:00:23	100.00%

图 14-67　百度搜索流量靠后的关键词

从 Google 带来流量的关键词前 16 位如图 14-68 所示。

网站使用 目标转化 浏览量：

访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
23,596 网站总额百分比：36.49%	1.42 网站平均数：1.55 (-8.64%)	00:00:44 网站平均数：00:01:01 (-26.78%)	94.73% 网站平均数：93.02% (1.83%)	84.45% 网站平均数：83.48% (1.16%)

	维度： 关键字	访问数	每次访问页数	平均网站停留时间	新访问次数百分比	跳出率
1.	亿技客	267	6.87	00:10:45	46.82%	24.34%
2.	e家网购物	110	1.00	00:00:00	100.00%	100.00%
3.	亿技客 蔡文卫	75	14.53	00:24:55	0.00%	16.00%
4.	蛋荞	40	1.02	> 00:00:00	97.50%	97.50%
5.	卡西欧指针mtp腕表	36	1.11	00:00:12	100.00%	94.44%
6.	荷泉(maxsun)8600gt留鼠笼	30	7.13	00:13:10	0.00%	10.00%
7.	亿客客	23	9.35	00:09:52	26.09%	0.00%
8.	site:www.yicike.com	20	11.15	00:11:23	15.00%	20.00%
9.	亿技购物网	18	1.06	> 00:00:00	5.56%	94.44%
10.	世川藤井	17	1.00	00:00:00	100.00%	100.00%
11.	p2222	16	1.12	00:00:38	93.75%	87.50%
12.	俏物俏语	16	3.88	00:03:37	68.75%	75.00%
13.	礼想家单身贵族咖啡泡茶兼容机	12	1.33	00:00:16	91.67%	75.00%
14.	老花镜价格	12	1.42	00:00:23	100.00%	75.00%
15.	318220-103	11	1.27	00:00:50	63.64%	63.64%
16.	lv男士手包	10	1.30	00:00:10	80.00%	70.00%

图 14-68　Google 搜索流量的前 16 个关键词

靠后一些的关键词如图 14-69 所示。

49.	t014.421.11.037.00	6	1.67	00:00:16	83.33%	50.00%
50.	乐梅网	6	1.17	00:00:13	83.33%	83.33%
51.	导购网	6	2.50	00:00:44	100.00%	50.00%
52.	小浪之家网上购物	6	1.17	00:01:52	100.00%	83.33%
53.	成人xiaoshuo	6	1.00	00:00:00	100.00%	100.00%
54.	海豚太阳镜	6	2.50	00:00:25	100.00%	83.33%
55.	电动键切机的价格	6	1.00	00:00:00	100.00%	100.00%
56.	石榴石的价格	6	1.00	00:00:00	100.00%	100.00%
57.	碎碳	6	1.00	00:00:00	100.00%	100.00%
58.	维维瘤盖瘤管养膜	6	1.00	00:00:00	100.00%	100.00%
59.	美的(midea)欧式抽油烟机cxw-200-dt17b	6	1.00	00:02:05	100.00%	83.33%
60.	西尔黑炎价格	6	1.17	00:00:06	100.00%	83.33%
61.	50度古井贡酒价格	5	1.00	00:00:00	100.00%	100.00%
62.	chanel项链价格	5	1.00	00:00:00	100.00%	100.00%
63.	crocs鞋价格	5	1.20	00:01:00	100.00%	80.00%
64.	dpr-1061	5	1.20	00:01:09	100.00%	80.00%
65.	e13kc1	5	1.00	00:00:00	100.00%	100.00%
66.	jy8009-55e	5	1.00	00:00:00	80.00%	100.00%
67.	ks-207d	5	1.20	00:00:58	100.00%	80.00%
68.	sex dvd	5	1.00	00:00:00	100.00%	100.00%
69.	t62缝纫锻	5	1.40	00:02:55	80.00%	60.00%
70.	tcl冰箱bc-92b	5	1.80	00:00:23	100.00%	60.00%

图 14-69　Google 搜索流量靠后的关键词

比较这几个抓图可以看到，除了涉及亿赐客品牌名称（当时叫亿枝客）的搜索外，带来流量的绝大部分是长尾关键词。最重要的网上购物、购物网站、网购等相关词没有排名也就没有流量。产品分类页面也很少带来流量，因为对应关键词（“数码产品”、“母婴用品”等）也很热门。

从 Google 和百度同时带来流量的关键词非常少，很少有页面在这两个搜索引擎上都有较好排名。

我也检查了几个关键词在百度和 Google 的具体排名，比如图 14-70 是“多普达 P2222”这个词在百度的排名。

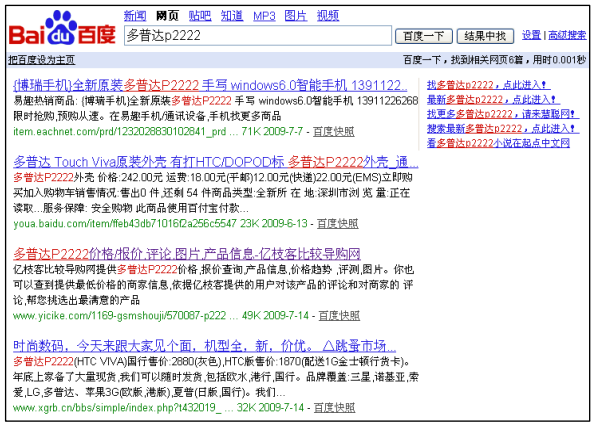


图 14-70　关键词“多普达 P2222”在百度的排名

图 14-71 是“motoe6 报价”这个词在百度的排名。

这两个带来流量靠前的词其实也都是很长尾的词，返回的页面很少，所以这种长尾

词排名还有进一步改进的可能。长尾关键词是需要投入时间、精力比较少，也比较现实的。由于各种原因，并不会因为有几个很强的竞争对手，就一定在长尾词上输给他们。不同的优化方式、不同的收录规模、不同的数据源等，使得再强大的竞争对手也一定有不能覆盖到的关键词。

图 14-71　关键词“motoe6 报价”在百度的排名

从图 14-72 可以看到，亿赐客网站在一个月时间内，有 44 202 个关键词带来过流量。这对一个有百万以上页面的网站来说是比较低的。下面有效收录页面的数字就更能说明这一点，如图 14-73 所示。

图 14-72　带来流量的关键词

网站流量工具所列出的所有搜索用户进入页面，可以理解为有效收录页面，因为只有这些页面才带来了搜索流量，其他页面虽然被收录，却一个流量都没有带来过。如图 14-73 所示，在一个月时间内，有 28 927 个页面带来过搜索流量，与整个网站百万页面级别相比，这显得比较低。

但另外一方面，相对于网站收录页面数在百度是 30 多万，在 Google 是低于 10 万，

这个有效收录页面比例就显得高多了。假定有效收录页面占所有收录页面之比固定，如果我们能大幅提高搜索引擎收录页面数，就意味着能大幅提高有效收录页面，也就是提高流量。

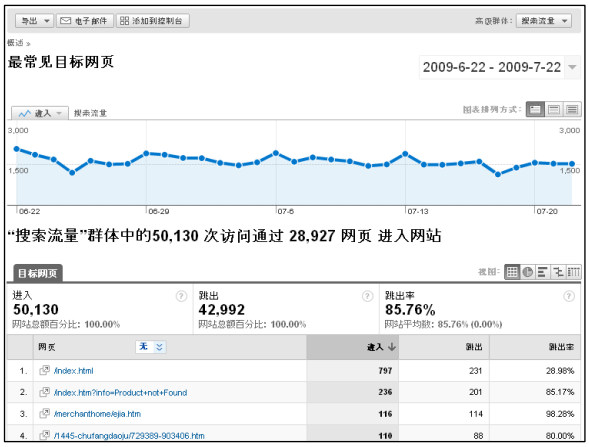


图 14-73　有效收录页面

所以我们可以得到一个结论，对亿赐客网站来说，要提高流量，关键在于提高收录率。从前面的统计可以看出，目前亿赐客网站收录页面数还太低，有很大改进空间。

当然，不同的网站必须具体问题具体分析，提高流量的关键可能有很大差别。这里所说的关键在于收录，只针对于亿赐客网站。

14.3.7　Google 网管工具数据和分析

Google 网管工具是一个非常强大、也非常有用的 SEO 工具。建议所有做 SEO 的人都要在网管工具注册自己的网站，能从中发现不少问题。下面就以亿赐客网站为例看看怎样从网管工具中发现对 SEO 有用的信息。

1．外链锚文字

图 14-74 显示的是 Google 数据库中指向亿赐客网站的外部链接所使用的锚文字（网管工具称为定位文字）。

外链锚文字是排名很重要的因素之一。从网管工具给出的数据，SEO 人员能够轻易地看出自己的目标关键词是否出现在外部链接锚文字中。如果这里列出的都是一些无关或次要的词汇，说明外部链接建设需要某种技巧上的改进。从抓图可以看到，亿赐客网站外链锚文字相对自然，但是针对主要关键词的外部链接比较少。

2．网站内容

图 14-75 显示的是 Google 在网站上最常见到的关键词。顺便提一下，很多人把“关

键词”称为“关键字”，Google 网管工具也是如此。我个人认为是很不恰当的用词。中文用户搜索主要是以“词”为基础的，而不是单个的“字”。

图 14-74　Google 网管工具显示的外链锚文字

图 14-75　Google 看到的网站内容和关键词

这个数据反映出网站内容是否真的与自己想优化的关键词相吻合。如果搜索引擎看到的关键词都是与你的目标关键词不相关的，网站主题与目标关键词有差距，自然很难获得好的排名。从抓图可以看到，亿赐客网站内容相关性还不错，诸如比较、产品、商家、商品、购物、搜索等都出现在最常见的关键词中。

3．速度影响抓取

图 14-76 显示的是 Google 抓取页面的统计，包括每天抓取的网页数量、字节数和下载页面所用的时间。

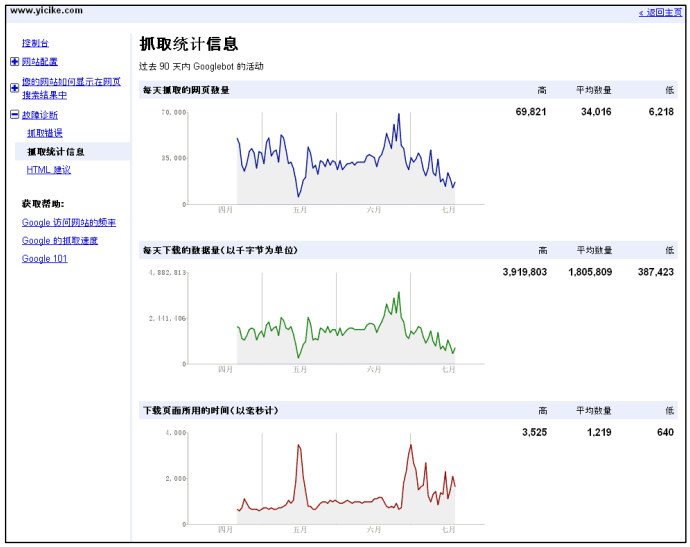


图 14-76　页面下载速度与抓取关系

从图 14-76 中我们可以明显看到一个现象，在 2009 年 5 月中旬的几天，当下载页面所用时间明显上升时，抓取的页面数量就急剧下降，也就是说网站下载速度直接影响了

页面抓取数量。这验证了前面提到过的，搜索引擎蜘蛛在一个特定网站上爬行、抓取的总时间是有限的。当然权重越高，这个限制值也就越大。亿赐客网站在 5 月份中旬的几天，因为某种原因服务器速度降低，每下载一个页面需要的时间大大增加，Google 蜘蛛爬行很少页面就用完了分配给亿赐客网站的总爬行时间，因此没有时间爬行更多页面。

如果读者看一下自己的网站，可能会发现很多网站下载所用时间与抓取网页数量不一定成反比，因为这些网站并没有用完 Google 给予这个网站的总爬行时间。

所以亿赐客网站需要尽量提高网站速度，包括优化数据库及减小页面文件，删除无关代码，这也很可能对提高收录率有帮助。

4．HTML 代码建议

Google 网管工具另一个非常重要的功能是 HTML 建议，Google 发现网站页面可能有问题时，会列在这里供站长参考。SEO 人员经常可以从这些建议中发现网站的重大问题，如图 14-77 所示。

图 14-77　网管工具 HTML 建议

下面就举几个亿赐客网站的例子。

（1）无标题标签

缺少标题标签是最简单也最容易更改的错误，站长自己却往往发现不了这个问题，如果能发现就不会不写标题标签了，一般都是因为疏忽大意。

从图 14-78 可以看到，出现这类问题的经常是“关于我们”、“联系我们”等容易被忽略的地方。

（2）重复内容页面

HTML 建议部分列出的重复元说明（也就是描述标签）、重复标题标签这两部分经常可以让站长发现网站上的重复内容或者网站架构上的问题。比如亿赐客网站就存在着几种原因造成的重复内容。

图 14-78　无标题标签页面

① 显示方式导致的重复标题及重复内容页面。

如图 14-79 所示，列在重复标题标签最前面的这几个页面，从 URL 就可以看出这是网站显示方式所造成的，也就是产品列表页面（也就是最末一级分类页面，通常是三级分类页面，有时也可能是二级分类页面）上按价格排列、按评论排列、按名称排列、按栅格方式显示、按列表方式显示等各种排列显示方法所造成的重复页面。

图 14-79　导航系统造成的重复标题标签和内容

网管工具列出的“护发乳护发素”分类页面，默认显示方式是图 14-80 所示的按栅格方式排列。

图 14-80　分类页面默认显示

如图 14-81 所示，用户点击按列表方式排列后，URL 变化，页面显示方式变化，但显示的内容是完全一样的，页面 Title 也一样。与此类似，用户点击右上角的按价格、按评价、按商品名称排序时都会生成不同的 URL，但页面内容重复。

图 14-81　分类页面按列表方式显示

② 技术问题导致的重复标题和重复内容页面。

我们再看重复标题部分列出的第三个页面，HOYA 多层镀膜产品页面，如图 14-82 所示。

图 14-82　技术问题造成的重复页面

列出的两个 URL 比较怪异，显然不是翻页或显示方式引起的。第一个页面的 URL 是：

http://www.yicike.com/2526-lvjing/613538-.htm

如图 14-83 所示。

第二个页面 URL 是：

http://www.yicike.com/2526-lvjing/613538-hoya49mm-uv-nhmc.htm

如图 14-84 所示。

图 14-83　页面 2526-lvjing/613538-.htm

图 14-84　页面 2526-lvjing/613538-hoya49mm-uv-nhmc.htm

第二个页面 URL 最后面的字符是从产品名称中的数字及英文字符提取的。这两个 URL 是同一个产品页面。

第一个 URL 去掉产品编号后面的横线：

http://www.yicike.com/2526-lvjing/613538.htm

显示的还是同一个产品页面。

更不妙的是，其实我们可以把产品编号之后的数字及字母部分改为任意其他字符，如：

http://www.yicike.com/2526-lvjing/613538-xxx.htm

如图 14-85 所示，网站系统还会正常显示同样产品的页面。也就是说只要产品系列

编号 613538 出现在 URL 中，不管后面加上什么字符，都显示同一个页面内容，而不是 404 错误。这是一个技术上的错误，可能导致无穷多个重复内容页面。由于某种原因，网站上确实出现了至少两个 URL 版本，还被搜索引擎收录了，也就是上面看到的，一个 URL 只出现产品编号，一个 URL 是产品编号加上产品名称中的数字及英文字符。这属于比较严重的技术问题。

图 14-85　产品编号后加任意字符

③ 考虑不周导致的重复标题及重复内容页面。

我们再来看一个例子，如图 14-86 所示。

图 14-86　考虑不周等造成的重复内容

“卫生纸”分类页面也出现了重复标题。第一个 URL 是正常的三级分类页面，如图 14-87 所示。

图 14-87　正常卫生纸分类页面

而第二个 URL 实际上是在正常分类页面上，用户点击左侧导航中“选择商家”下面的商家链接后所到达的页面，比如点击“1 号店”将来到如图 14-88 所示的页面。

问题在于，选择了商家 1 号店之后，页面标题标签、面包屑导航、显示方式下面的说明文字都没有任何变化，与卫生纸分类首页完全相同，用户很难感觉到已经将产品限制在某个商家。这既对用户不友好，用户无法从页面看出这实际上是一号店卫生纸产品列表，对搜索引擎也不友好，搜索引擎同样无法分辨这两个页面的区别。

图 14-88　点击“1 号店”后达到页面

④ 同一产品出现在不同分类导致的重复标题及重复内容页面。

我们再看下面“兰亭序 VCD”产品页面。从列出的重复页面 URL 我们就可以看出，由于某种原因，“兰亭序 VCD”这个产品出现在两个不同分类中，分别是分类编号 1685 和 1687。点击这两个 URL，确实打开的是同一个产品页面，只不过 URL 不同，如图 14-89、图 14-90 所示。

图 14-89　“兰亭序 VCD”在分类 1685 中

图 14-90 “兰亭序 VCD”在分类 1687 中

从 URL 中的拼音看，这个产品似乎更应该属于编号为 1685 的影视分类。如果我们直接访问 1687 分类页面，会看到这实际上是太阳伞分类列表，如图 14-91 所示。

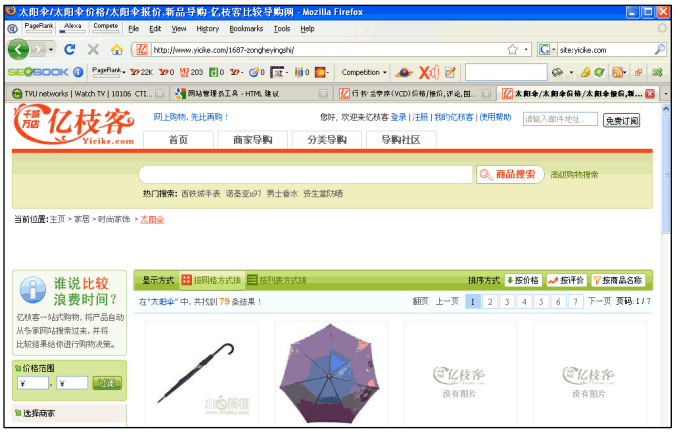


图 14-91 编号为 1687 的分类

不知道为什么“兰亭序 VCD”这个产品会出现在这两个分类目录中。

从上面的几个例子可以看到，在 Google 网管工具中能发现的问题，常常是观察网站本身很难注意到的。对稍微有些规模的网站，你不太可能访问和研究网站的大部分页面，看链接和 URL 是否有问题。查看搜索引擎收录的所有页面 URL，也只能看有限页数而已。对这种几十万几百万页面的网站来说，很难通过人工观察网站发现这种隐藏的重复内容页面和技术问题。这就是 Google 网管工具的重要意义所在。

14.4 关键词研究

针对亿赐客网站情况，我把整个网站分为 6 种页面，分别研究需要优化的关键词。

这 6 种页面是：

- 首页
- 分类页面
- 商家页面
- 品牌页面
- 产品页面
- 搜索页面

14.4.1 首页

首页是整个网站权重最高、排名能力最强的页面，通常把最热门的关键词放在首页上。这里有两个选择，一是针对比较购物类的关键词，像“比较购物”、“购物搜索”、“导购网”等，二是针对更通用的购物关键词，比如“网上购物”、“购物网站”、“网购”、“网上商城”等。

图 14-92 是 Google 趋势所显示的这两类关键词热门程度，“网上购物”这类词的搜索数量远远大于“比较购物”类。

图 14-92　两类关键词搜索次数对比

从时间上看，2006 年以后两类关键词都呈上升趋势，但是在 2008 年下半年国际金融风暴之后，搜索次数有所下降，说明整体经济形势也影响了用户在网上购物的意愿。

百度指数也显示，“网上购物”和“比较购物”这两类关键词搜索数量差距很大。“网上购物”指数在 2000 以上，而“比较购物”指数只有两位数。所以就算“比较购物”这类词排到第一，能带来的流量也非常有限，甚至可以说，从搜索流量角度看，这种词对亿赐客网站没有意义。

鉴于这种情况，我觉得亿赐客网站应该把首页的目标关键词定在“网上购物”这种搜索量更大的关键词上。

然后使用 Google 关键词工具，进一步挖掘相关关键词搜索情况。表 14-3 是从 Google 关键词工具得出的一部分搜索量数据。

表 14-3　网上购物类关键词扩展

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
[网上购物]	1	301 000	301 000
[背背佳]	0.8	40 500	135 000
[易趣]	0.86	90 500	135 000
[安利]	0.93	110 000	110 000
[购物网站]	0.93	90 500	90 500
[宜家]	0.8	90 500	90 500
[红孩子]	0.73	135 000	74 000
[团购]	0.86	201 000	74 000
[网上商城]	0.93	18 100	49 500
[淘宝商城]	0.66	49 500	49 500
[网上书店]	0.8	22 200	40 500
[亚马逊]	0.73	60 500	40 500
[日本代购]	0.73	–1	33 100
[九阳豆浆机]	0.66	22 200	33 100
[购物网]	0.93	33 100	33 100
[网购]	0.8	246 000	33 100
[佐丹奴]	0.73	18 100	27 100
[易趣网]	0.8	22 200	27 100
[橡果国际]	0.66	18 100	22 200
[豆浆机]	0.86	14 800	14 800
[网上买东西]	0.6	720	12 100
[qq 商城]	0.46	–1	9900
[网上购物网站]	0.8	5400	9900
[购物论坛]	0.73	–1	8100
[代购]	0.8	–1	8100
[导购网]	0.53	–1	6600
[网购手机]	0.6	–1	4400
[易趣购物网]	0.8	2400	4400
[美国代购]	0.6	–1	3600
[代购网]	0.66	–1	2900
[手机导购]	0.93	2400	2900
[网上超市]	0.66	2900	2400
[手机商城]	0.73	6600	2400

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
[大拿网]	0.4	–1	1600
[香港网上购物]	0.6	–1	1600
[美国购物网]	0.53	–1	1600
[网上商店系统]	0.6	880	1300
[比较购物]	0.53	–1	1000
[聪明点]	0.46	–1	1000
[网络商店]	0.6	–1	1000
[日本购物网站]	0.6	–1	880
[美国购物网站]	0.46	–1	880
[网上店铺]	0.66	1000	880
[大拿]	0.26	–1	720
[手机网购]	0.53	–1	720
[网上代购]	0.6	–1	590
[香水网上]	0.33	–1	590
[手机超市]	0.46	390	590
[购物搜索]	0.6	1900	590
[网上商场]	0.6	–1	480
[网上手机商城]	0.53	210	480
[淘宝网上购物]	0.6	1000	480
[淘宝导购网]	0.2	–1	390
[在线购买]		170	320
[网购图书]	0.26	–1	260
[在线购物网站]	0.66	140	260
[网上购物商店]	0.66	210	260
[网上购物超市]	0.53	110	210

删除其中不太相关的词以及搜索量很小的词之后，挑出表 14-4 中显示的最重要的关键词。

表 14-4　主要关键词搜索次数及竞争程度

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量	搜索结果数	前两页强域名数	前两页域名首页数	allintitle 指令数字	广告数
[导购网]	0.53	–1	6600	23 600 000	1	19	6 110 000	1

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量	搜索结果数	前两页强域名数	前两页域名首页数	allintitle 指令数字	广告数
[网上买东西]	0.6	720	12 100	10 800 000	12	1	18400	3
[网上超市]	0.66	2900	2400	14 900 000	3	15	3 400 000	1
[网上购物网站]	0.8	5400	9900	12 900 000	7	18	95 400	3
[网上商城]	0.93	18 100	49 500	22 100 000	11	16	19 600 000	8
[购物网站]	0.93	90 500	90 500	13 200 000	6	17	36 000 000	6
[网购]	0.8	246 000	33 100	19 700 000	5	19	4 080 000	6
[网上购物]	1	301 000	301 000	21 000 000	7	16	21 900 000	8
[购物搜索]	0.6	1900	590	13 000 000	5（4 个竞争对手）	16	55 700 000	1
[在线购买]		170	320	13 100 000	4	3	2 690 000	0

除了 Google 关键词工具给出的广告客户竞争程度、本地搜索量、全球搜索量之外，还列出了另外几个数据。

- 搜索结果数：搜索结果数越大，参与竞争的页面数越多，要获得好的排名也越困难。
- 前两页强域名数：排在前两页的强域名越多，竞争也越厉害。这里所说的强域名不一定准确，比较个人化。我所列出的强域名指的是我个人听说过的网站，诸如大的门户、著名电子商务网站等。
- 前两页域名首页数：如果排在前面的都是域名首页，说明竞争程度高，网站内页很少有机会排到前面。
- Allintitle 指令数字：也就是页面标题标签中出现相应关键词的搜索结果数。这个数字越大，说明针对这个关键词优化的页面数越多，也就是竞争越厉害。
- 广告数：显然广告越多，竞争越强。

观察表 14-4，我觉得下面三个词可以列为亿赐客首页的目标关键词：“网上购物”，“网购”及“购物网站”。而“网上购物”和“购物网站”连在一起，又可以生成“网上购物网站”这个搜索数也还不错的关键词。

表格中的搜索数量我使用了广泛匹配，为的是从总体上看关键词的热门程度，而不是预估流量，所以数字比完全匹配要高出 3～4 倍。比如“网上购物”这个词，广泛匹配每个月有 30 万搜索量，使用完整匹配的话，每个月只有 10 万左右的搜索量，每天 3000 多次搜索。

从搜索量看，就算“网上购物”、“购物网站”这类搜索量比较大的词排到前面，带来的流量也不会很高。图 14-93 显示的是 Google Adwords 点击量预估工具列出的数字。

“网上购物”这个词，如果广告商链接排在 1～3 位，每天所能带来的流量也只不过是 50 个左右。左侧自然排名的点击量通常会比右侧广告多 5 倍以上，也就是说自然搜索流量只能带来两三百个 IP 而已。考虑到 Google 和百度各自在中国的市场占有率，网上购物、购物网站这类词即使能排到百度和 Google 第一（这几乎是不可能的），带来的搜索流量也不会超过几千个。从这个意义上来说，要增加网站的流量还是要靠数量巨大的长尾关键词和产品页面，热门关键词能起到的作用有限。

点击量估算工具

修改设置 | 以 csv 格式下载

所有估算值仅作为参考指南提供。它们是根据整个系统内的平均值计算的。您的实际费用和广告排名可能有所不同。要查看根据您关键字的效果历史记录计算的估算值，请使用相应广告组中的点击量估算工具。了解详情

平均每次点击费用: ¥2.61 (最高每次点击费用为 ¥11.54)
估算的每日点击次数: 764 - 948 (每日预算为 ¥2,600.00)

估算值取决于您的出价以及地理定位选择。因为点击量估算工具不会考虑您的每日预算，您的广告获得的点击次数可能低于估算值。

最高每次点击费用: 每日预算 获得新的估算值

关键字	搜索量	估算平均每次点击费用	估算的广告排名	估算的每天点击量	估算的每天费用
搜索网络总计		¥1.93 - ¥2.78	1 - 3	764 - 948	¥1,500 - ¥2,600
比较购物		数据不足无法算出估算值。			
充值卡		¥0.57 - ¥0.95	1 - 3	8 - 10	¥10
导购网		¥0.72 - ¥1.20	1 - 3	0	¥10
购物		¥1.19 - ¥1.89	1 - 3	179 - 227	¥300 - ¥500
购物网站		¥1.88 - ¥2.85	1 - 3	7 - 9	¥20 - ¥30
化妆品		¥2.06 - ¥2.86	1 - 3	137 - 174	¥300 - ¥500
数码相机		¥2.25 - ¥3.46	1 - 3	96	¥300 - ¥400
网购		¥1.09 - ¥1.90	1 - 3	83 - 105	¥100 - ¥200
网购书		¥2.90 - ¥4.35	1 - 3	5 - 7	¥20 - ¥40
网购物		¥0.67 - ¥1.23	1 - 3	43 - 54	¥30 - ¥70
网上购物		¥1.84 - ¥2.71	1 - 3	45 - 62	¥90 - ¥200
网上商城		¥3.47 - ¥4.52	1 - 3	16 - 21	¥60 - ¥100
鲜花		¥2.87 - ¥3.73	1 - 3	143 - 182	¥500 - ¥700
搜索网络总计		¥1.93 - ¥2.78	1 - 3	764 - 948	¥1,500 - ¥2,600

图 14-93　点击量预估

再来看一下这几个主要关键词目前是哪些网站排在前面，亿赐客是否有机会获得好的排名。2009 年 7 月，在 Google 搜索“网上购物”，排在前面的 10 个页面是：

（1）www.taobao.com/

（2）www.taobao.com/index_n.php

（3）www.hao123.com/netbuy.htm

（4）www.amazon.cn/

（5）www.amazon.cn/b/63153

（6）www.dangdang.com/

（7）www.265.com/Wangshang_Gouwu/

（8）www.zhigou.com/

（9）bbs.egou.com/

（10）www.smarter.com.cn/

搜索“购物网站”排在前面的 10 名是：

（1）www.hao123.com/netbuy.htm

（2）www.taobao.com/

（3）www.1b2g.com/shop.php

（4）www.eachnet.com/

（5）www.amazon.cn/

（6）site.baidu.com/list/33wangshanggouwu.htm

（7）www.usashopcn.com/

（8）www.dangdang.com/

（9）www.7shop24.com/

（10）www.wooha.com/

这两个比较热门的关键词，排在前面的多是淘宝、当当这种巨无霸级别的网站，要跻身第一页难度很大，暂时可以不用考虑了。

所以就首页来说，我们的策略是顺其自然，瞄准应该瞄准的，但短期不寄希望于能排到前面。

在首页的文字内容上，需要参考与“网购”语义相关的词汇。使用 Google Sets 工具可以生成一系列语义相关的词，供文案写作时参考。比如 Google Sets 工具提示与“网购”相关的词包括：网上购物，购物网站，购物网，团购，代购，淘宝，店铺，商品，价格，旺旺，服务，图片，购买，钻石，朋友等。

14.4.2 分类页面

网站产品分类已经固定，关键词研究所能做的主要不是寻找主关键词，而是找出每一个分类应该使用哪种组合描述方法，title、heading、页面文字等都需要按关键词搜索次数安排。我们分类时所使用的词与用户真正搜索的词可能并不相同，需要进行基本的关键词研究。

网站有 18 个一级分类，每一个分类我都用 Google 关键词工具进行了分析，列出搜索次数比较多的备选关键词。以“家电”分类为例，表 14-5 是 Google 关键词工具显示的相关关键词搜索次数。

表 14-5 “家电”分类关键词搜索次数

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
家电	0.93	2 740 000	1 500 000
家用电器	0.86	1 220 000	450 000

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
家电团购	0.66	2900	2900
宁波家电	0.26	3600	2900
品牌家电	0.46	27 100	18 100
小家电	0.86	110 000	246 000
家电网	0.6	60 500	60 500
促销 家电	0	2900	4400
家用电器 电视	0	–1	9900
团购家电	0.26	2900	2900
三星家电	0.53	1900	1300
家用电器产品	0.33	–1	1900
松下家电	0.33	–1	1000
武汉家电	0.33	–1	2900
西门子家电	0.6	4400	2900
家电厂	0.26	–1	1300
家电电视	0	–1	33 100
工贸家电	0.53	–1	2900
广东家电	0.33	–1	1900
电器家电	0.33	–1	3600
上海永乐家电	0.53	–1	2900
家电公司	0.53	–1	6600
家电杭州	0.2	6600	5400
选购家电	0	390	2400
美的家电	0.6	14 800	9900
中国家电	0.6	40 500	27 100
中国家电网	0.6	9900	9900
家电 宁波	0	3600	2900
永乐 家电	0.53	14 800	14 800
永乐家电 上海	0	–1	2900
家电代理	0.53	3600	5400
家电报价	0.6	22 200	14 800
小家电网	0.6	5400	2400
家电厂家	0.4	–1	880
慈溪家电	0.4	–1	1000

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
九阳家电	0.4	–1	1900
佛山家电	0.13	–1	1000
农村家电	0	–1	1000
南京家电	0.2	–1	1900
南昌家电	0	–1	6600
大连家电	0	–1	880
天津家电	0.13	–1	1600
家电交易	0	–1	1300
家电加盟	0.46	–1	1000
家电商场	0.46	–1	2400
家电展	0.53	1900	1900
小家电批发市场	0.4	–1	1300
成都家电	0.33	2900	2900
沈阳家电	0	–1	1600
西安家电	0	–1	1900
郑州家电	0	–1	1000
重庆家电	0	–1	1900
长沙家电	0	–1	1000
长虹家电	0.4	1900	1600
青岛家电	0.4	–1	1000

删除不太相关的一些词，如“中国家电”、“家电网”、“家电公司”后，搜索次数最多的是“家电”、“家用电器”、“小家电”、“品牌家电”、“家电报价”这几个词。

另外，还可以观察到有两类搜索词很常见，一是“地名+家电”，比如南昌家电，宁波家电，武汉家电，成都家电，广东家电等。另一种是“品牌名+家电”，如永乐家电，美的家电，西门子家电，长虹家电，三星家电等。这是非常有价值的信息。前面关键词研究章节提到过，关键词研究是网站内容策划、扩展的重要来源和依据之一，下面网站优化建议部分我们再详细讨论怎样根据这些观察到的关键词扩展内容、更改网站架构。

其他分类原理完全一致，这里就不再详细提供列表，只给出结论，也就是每个一级分类搜索次数比较高的备选关键词。

- 图书：图书批发，图书购买，图书网，图书音像，特价图书，网上书店。
- 娱乐：最新游戏，最新音乐，电影娱乐。
- 服装鞋帽：品牌服装，韩国服装，流行服装，时尚服装。

- 珠宝饰品：珠宝饰品，时尚服饰，流行服饰，珠宝首饰，韩国服饰，休闲服饰。
- 礼品鲜花：鲜花速递，网上订花，礼品网，礼品公司，商务礼品，礼物，工艺品，地名+鲜花，地名+礼品。
- 手机通讯：手机软件，手机配件，手机报价，智能手机，最新手机，品牌名+手机，地名+手机。
- 电脑：笔记本电脑报价，电脑价格，手提电脑，特价电脑，购买电脑。
- 数码产品：数码产品报价，数码摄像机，数码商城。
- 家电：家用电器，家电报价 ，小家电，品牌家电，地名+家电，品牌名+家电。
- 化妆护肤：化妆品，日本化妆品，化妆品价格，品牌化妆品，韩国化妆品。
- 汽车用品：汽车用品，汽车装饰用品，汽车美容，汽车零部件，汽车音响，地名+汽车用品。
- 家居园艺：家具用品，家具饰品，家具装饰，宜家家居，时尚家居，家具建材，园艺用品，智能家居。
- 母婴用品：母婴用品店，宝宝用品，婴儿用品，儿童用品，母婴用品批发。
- 运动户外：户外运动，运动器材，运动服装，运动休闲。
- 健康医药：健康医药，医疗器械，健康饮食，医疗保健，两性健康。
- 办公用品：办公用品，办公家具，办公文具，公司办公用品，办公设备，地名+办公用品。
- 食品饮料：食品饮料，食品网，绿色食品，休闲食品，进口食品，地名+食品。
- 玩具宠物：宠物用品，模型玩具，毛绒玩具。

网站上还有更大量的二级和三级分类，原则上也应该使用工具检查关键词搜索次数，然后列出搜索最多的 2～3 个关键词，作为相应分类页面的最主要目标关键词。二级和三级分类加起来数量上千，工作量很大，这部分只能留给网站的 SEO 人员去慢慢进行了。

图 14-94 是 Google 趋势显示的几个分类关键词搜索量，其趋势与首页上的“网上购物”等词大致相同。

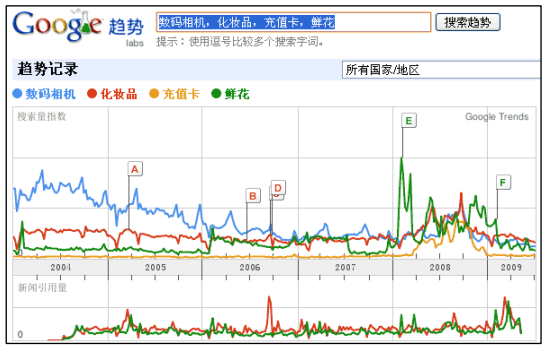


图 14-94　分类页面关键词趋势

再来看一下分类页面排名难度。图 14-95 显示的是在 Google 搜索“数码相机”的结果页面及 SEO for Firefox 显示的各页面概况。

图 14-96 显示的是化妆品分类情况。

从图中可以看到，排在这两个一级分类关键词前面的都是很著名的门户或电子商务网站，尤其是数码相机分类，竞争更为激烈，都是大门户的频道页面。要想排进前 10 名，可能性极低。

图 14-95 “数码相机”搜索结果页面

图 14-96 “化妆品”搜索结果页面

图 14-97 显示的是搜索“手机电池”这个二级分类名称结果。

排在前面的虽然还基本上是大网站，如中关村在线、阿里巴巴等，区别在于出现了一些不是这么强的 B2C 网站。另一个更重要的数字是，排名页面的雅虎链接数相比前面两个一级分类而言，下降了很多，排在前面的一些内页外链只有几十或几百个。二级分类这个级别已经开始出现能获得排名的希望。

图 14-98 显示的是“珍珠粉”这个三级分类名称搜索结果。

图 14-99 显示“豆浆机”这个三级分类名称的情况。

三级分类形势更加乐观，出现了不少内页，雅虎链接数有几个是个位数，说明亿赐客网站三级分类对应关键词获得排名的可能性比较大。除了首先关注长尾流量，三级分类名称可以作为较热门关键词的突破口。当然，前提是域名权重足够高，而且三级分类页面也有一定数量的外部链接。

图 14-97 “手机电池”搜索结果页面

图 14-98 “珍珠粉”搜索结果页面

图 14-99 “豆浆机”搜索结果页面

14.4.3 商家页面

亿赐客网站上每个商家都有一个单独页面介绍商家情况。这些商家页面除了商家名称本身，还可以优化哪些关键词呢？图 14-100 至图 14-103 是百度的相关搜索，显示在搜索“当当”、“京东商城”、“凡客诚品”、“红孩子”等商家名称时，用户还搜索了与之相关的哪些关键词。

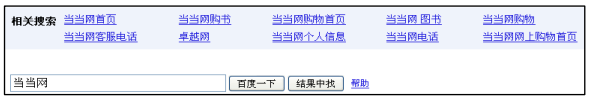


图 14-100 “当当网”相关搜索

图 14-101 “京东商城”相关搜索

图 14-102 “凡客诚品”相关搜索

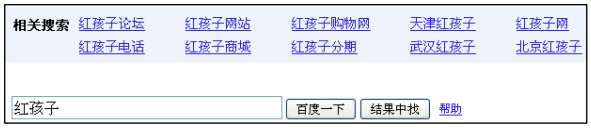


图 14-103 “红孩子”相关搜索

从这四个抓图可以看到几个经常出现又能适用于所有商家的扩展词：网站、优惠券、客服电话、地址、怎么样、评价等。这几个词与商家名称搭配组合，可以生成不少关键词，商家页面就是针对这类关键词优化的最合适页面。

像这种扩展词 SEO 人员自己也可以想到一些，但毕竟不如搜索引擎告诉你用户真正在搜索什么更准确。这类相关扩展词使用 Google 关键词工具也可以得到类似结果。使用哪个工具并不重要，重要的在于从不同搜索关键词中找到规律，运用在自己的网站上。

诸如“图书”、“女装”这种扩展词只能用于某些商家，SEO 人员有时间人工调整页面 Title、正文文字时可以考虑使用。用于模板式优化的只能是前面列出的适用于所有商家的词。

14.4.4 品牌页面

与商家页面类似，可以给每个品牌都建立一个页面，优化相关关键词和组合。

2009 年 7 月时亿赐客网站并没有品牌页面，但其实与品牌相关的搜索数量不小，这是关键词研究决定网站内容的又一个例子。

除了品牌名称之外，品牌页面还可以融入哪些词呢？使用 Google 关键词工具可以观察到一些常见的、与品牌名称相关的词语，其原理和上面商家页面使用百度相关搜索一样。

表 14-6 至表 14-8 分别是与索尼、惠普、华硕这几个品牌名称相关的搜索词。其中包含不少只和这个品牌有关、而不具普遍性的词，比如索尼电脑、索尼随身听、惠普打印机等。电脑、随身听、打印机这些词不能普遍适用于所有品牌名称，只能用于人工调整。

表 14-6 “索尼”相关搜索

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
索尼摄像机	0.66	74 000	74 000
索尼电脑	0.66	74 000	49 500
索尼	0.86	2 240 000	1 830 000
索尼随身听	0.4	2900	1900
索尼有限公司	0.33	–1	4400
索尼中国	0.53	9900	9900
东芝 索尼	0.2	–1	260
投影机 索尼	0	8100	5400
索尼 评测	0.2	–1	18 100
dv 索尼	0	5400	4400
md 索尼	0.2	2400	1300
摄像机 索尼	0.2	74 000	74 000
电脑 索尼	0	74 000	49 500
笔记本 索尼	0.2	135 000	110 000
索尼 产品	0.26	–1	–1
索尼 价格	0.26	–1	22200
索尼 评价	0.26	–1	1600
索尼 测评	0.26	–1	3600
耳机 索尼	0.13	18100	14800
索尼 f828	0.33	590	880
索尼 15e	0.26	–1	46
psp 索尼	0	27 100	27 100
上海索尼	0.33	8100	8100
数码索尼	0.46	368 000	368 000

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
索尼 20e	0.26	–1	28
索尼 dcr	0.33	–1	12100
索尼 dv	0.46	5400	4400
索尼 f717	0.4	1000	1000
索尼 hc15e	0.26	–1	110
索尼 md	0.33	2400	1300
索尼 psp	0.6	27 100	27 100
索尼上海	0.2	8100	8100
索尼中	0.2	–1	480
索尼中国网站	0.2	–1	1000
索尼中文	0.2	–1	1900
索尼中文网站	0.2	–1	320
索尼产品	0.33	6600	4400
索尼介绍	0.26	–1	1300
索尼官方	0.33	–1	14 800
索尼官方网站	0.46	12100	12 100
索尼报价	0.33	–1	33 100
索尼数码	0.66	368 000	368 000
索尼笔记本	0.8	135 000	110 000
索尼网站	0.4	14 800	14 800
中國索尼	0	–1	73
索尼中國	0.2	–1	73
索尼爱立	0.4	1000	880
索尼说明书	0	–1	8100
索尼配件	0.33	–1	1900
索尼集团	0.26	–1	590
索尼驱动	0.26	–1	8100
索尼公司	0.46	14 800	14 800
索尼投影机	0.73	8100	5400
索尼首页	0.2	–1	170
中国索尼	0	9900	9900
索尼耳机	0.53	18 100	14 800
索 尼 中 国	0	–1	22

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
索尼电池	0.46	14 800	12 100
数码摄像机 索尼	0.13	33 100	33 100
电视机 索尼	0	8100	8100
收音机 索尼	0	5400	9900
彩电 索尼	0	4400	4400
ibm 索尼	0	–1	73
索尼 参数	0	–1	2400
电池 索尼	0	14800	12 100
索尼 中国 有限公司	0.4	–1	–1
索尼 爱立 信	0.46	–1	–1

表 14-7 “惠普”相关搜索

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
惠普	0.86	1 830 000	1 500 000
惠普 hp pavilion	0	–1	590
惠普 中国	0.33	–1	–1
上海 惠普	0.26	–1	–1
惠普 评测	0.2	–1	14 800
惠普 6315	0.26	–1	–1
招聘 惠普	0	4400	8100
惠普 photosmart	0.2	–1	1600
惠普 打印机	0.6	–1	–1
联想 惠普	0	–1	720
绘图仪 惠普	0	4400	2400
惠普 服务器	0.4	–1	–1
惠普 制造商	0.2	–1	–1
惠普 笔记本	0.53	–1	–1
惠普 计算机 产品 上海 有限公司	0	–1	–1
惠普 博 客	0.2	–1	–1
惠普 音响	0.26	–1	–1
显示器 惠普	0	–1	1900
hp 惠普	0.46	–1	49 500
惠普电脑	0.8	201 000	165 000

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
惠普畅游	0.2	–1	720
惠普市场	0.2	–1	480
爱惠普	0.4	1900	1900
惠普墨	0.2	–1	2400
上海惠普	0.4	12 100	12 100
惠普 pavilion	0.33	–1	5400
惠普机	0.13	–1	14 800
惠普性能	0.2	–1	390
惠普科技	0.26	–1	880
惠普维修	0.6	33 100	27 100
墨盒 惠普	0	27 100	18 100
惠普专卖	0.53	3600	3600
惠普 nx9040	0.2	–1	46
惠普 1940	0.2	–1	91
惠普打印机	0.8	135 000	110 000
惠普 cn	0	–1	1000
惠普墨盒	0.73	27100	18 100
惠普 pda	0.53	1900	2900
惠普 nc6000	0.26	–1	480
惠普公司	0.53	14 800	14 800
惠普激光打印机	0.66	22 200	12 100
电脑惠普	0.2	201 000	165 000
惠普 2210	0.26	–1	210
惠普 m2000	0.2	–1	110
惠普驱动	0.4	74000	74 000
惠普报价	0.33	–1	18 100
惠普 评价	0.13	–1	1000
惠普 ibm	0.13	–1	260
惠普笔记本	0.8	368 000	301 000
惠普价格	0	–1	12 100
激光打印机 惠普	0	22 200	12 100
惠普 测评	0.13	–1	880
惠普系列	0.2	–1	2900

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
惠普 3538	0.2	–1	210
惠而普	0.2	–1	880
pda 惠普	0	1900	2900
惠普产品	0.33	–1	880
惠普 评论	0.2	–1	390
惠普 1010	0.26	–1	1000
惠普 dv1000	0.26	–1	260
惠普服务	0.4	18100	14 800

表 14-8 “华硕”相关搜索

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
华硕	0.8	1 220 000	1 220 000
机箱 华硕	0	–1	1000
华硕 评测	0.13	–1	18 100
手机 华硕	0	–1	8100
显卡 华硕	0	–1	18 100
超频 华硕	0	–1	2400
光驱 华硕	0	–1	1900
cpu 华硕	0	–1	1600
intel 华硕	0	–1	260
华硕 测试	0.2	–1	320
华硕 参数	0.13	–1	1600
华硕 m2npv vm	0.2	–1	–1
华硕 a8 测评	0.2	–1	–1
华硕电脑	0.73	135 000	110 000
华硕 s200	0.33	480	720
华硕 a730	0.26	110	210
华硕 a620+	0.26	–1	22
华硕 评论	0.2	–1	390
华硕笔记本	0.73	246 000	246 000
华硕报价	0.26	27100	27100
华硕 测评	0.2	–1	1600
电脑华硕	0	135 000	110 000
华硕 m2400	0.26	–1	1000

续表

关 键 字	广告客户竞争程度	本地搜索量：6 月	全球每月搜索量
华硕 价格	0.2	–1	8100
华硕系列	0.26	–1	3600
华硕 m6n	0.26	–1	36
华硕 s200n	0.26	–1	140
华硕 s300	0.26	–1	73
华硕 s300n	0.2	–1	91
华硕促销	0.2	–1	260
华硕驱动	0.26	–1	49 500
华硕 m5	0.2	–1	170
华硕程序	0.2	–1	1600

那些能普遍适用于所有品牌的扩展词才能在品牌页面上自动生成文字内容。观察一下上面三个列表，产品、评价、评测、服务、专卖店、网站、公司这些词经常会与品牌名称一起搜索，而且可以适用于所有品牌。这些词加上品牌名称，就可以作为品牌页面的目标关键词。

14.4.5 产品页面

显然，产品名称是产品页面首要目标关键词。还有哪些扩展词可以适用于所有产品，自动优化进所有产品页面中？图 14-104 至图 14-111 是在百度搜索不同产品名称时得到的相关搜索。

图 14-104 “九阳豆浆机”相关搜索

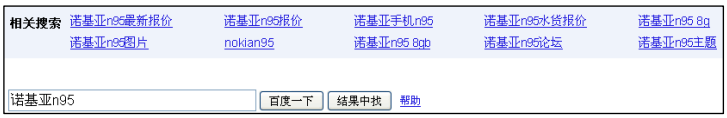


图 14-105 “诺基亚 n95”相关搜索

图 14-106 “安娜苏许愿精灵香水”相关搜索

同样，仔细观察这些以产品名称为搜索词的相关搜索，经常出现产品名称加上价格、

报价、怎么样、好吗、真假这几个扩展词，这些词也不局限于特定产品，可以加在任何产品名称后。这些词都有真实的用户在搜索，SEO 人员自己很难靠想象了解全面。比如“九阳豆浆机怎么样”这种词，很少有电子商务网站专门做页面来进行优化，发现了这些带有一定规律性的扩展词，就可以有意识地在页面上进行优化。

图 14-107 “雅顿绿茶女士香水”相关搜索

图 14-108 “ZIPPO 经典铭系列”相关搜索

图 14-109 “LG 双开门冰箱”相关搜索

图 14-110 “飞利浦剃须刀 HQ6073”相关搜索

图 14-111 “先锋家庭影院”相关搜索

14.4.6 搜索页面

搜索页面无法进行特定的关键词研究，不过前面的关键词研究已经提示我们，有些关键词很难放在上面 5 类页面上，却可以简单地作为搜索页面的目标关键词。

比如前面提到的“宁波家电”、“广东家电”这类词，要在网站主体分类结构中进行优化比较困难，很难有逻辑性地将其放入哪个分类。而将这些词做成搜索页面则顺理成章。搜索页面和 tags 页面类似，页面之间没有从属关系，也就无须考虑结构关系，只要在其他页面出现爬行和抓取入口（搜索页面的链接）就可以。这类词其实不少，比如地名+办公用品，地名+食品，地名+汽车用品等，而且搜索量不小。

细心的读者可能发现，这个案例中所讲的关键词研究，与前面关键词研究章节讨论

的一般性方法有些区别。通常一般性的关键词研究是要找到搜索次数比较多、竞争比较小的关键词，这样得到好排名的可能性比较大。而亿赐客网站关键词研究所做的并不是这样。可以说，我们基本上不太考虑能获得好排名的可能性，而是直接把目标放在了搜索次数比较多的关键词上，不管最后能不能排上去。这样做有两个原因：

（1）亿赐客这样的网站关键词包罗万象，由产品分类决定，几乎已经无法改变，或者说产品的分类不是 SEO 人员所能确定的，所以 SEO 人员明知某些关键词难度大，也不能改变。

（2）热门关键词至少在网站优化的头一两年不是重点，能做到什么程度就做到什么程度，带来流量的重点是长尾关键词和产品页面。

这个案例也可以说明，SEO 必须具体问题具体分析，没有一个适合于所有网站的金科玉律。

结合上面的关键词分析和竞争对手分析，可以得到几点结论。

- 首页和一级、二级分类页面目标关键词要得到好排名的可能性很小，我们就顺其自然，做好内部优化，近期不寄望有排名。
- 三级分类关键词有获得排名的可能性。除了页面本身的优化之外，友情链接交换可以从三级分类页面开始。
- 主要竞争对手有 5～6 个，其网站实力、历史都比亿赐客要强，但还没有一个占绝对性、压倒性优势的中文比较购物网站。
- 2009 年第一梯队的比较购物网站流量日 IP 在 5 万左右，这也就是亿赐客网站 SEO 的目标：在一到两年内搜索流量达到 5 万日 IP，进入第一竞争梯队。
- 提高搜索流量的关键在于长尾词和产品页面。
- 获得长尾流量的关键在于网站收录，这就需要在网站架构、页面内容的扩充上下一番工夫。
- 几个主要竞争对手应该是经过专业 SEO 优化的，尤其是聪明点和智购。

14.5 亿赐客网站优化建议

下面是 2009 年 10 月提供给亿赐客团队的优化建议。为保持原意和真实性，除了错别字、病句之类的明显错误，没有做什么修饰，读者在这里看到的基本上就是亿赐客团队收到的。因此，请读者包涵文字的粗糙。

报告里比较详细地写了怎样修改，有的地方写了为什么，有的地方没有写。没有写为什么的，请读者参考本书前面章节，都可以找到答案。

读者看到本书时，亿赐客网站必然已经和我当初诊断时不同了。我尽量留下原始 URL、代码和抓图等资料，以使读者了解诊断的原始对象。即使这样，下面的报告还是相当烦琐、枯燥的，有的地方可能不容易明白。真正能静下心来看完、看明白这个报告的

读者，相信会有很大收获，尤其是没有 SEO 实战经验，面对网站不知道如何下手的新手。

图 14-112 至图 14-121 是 2009 年 7 月我诊断网站时的抓图，读者看下面诊断及建议时可能需要经常参考。

图 14-112　亿赖客网站首页

图 14-113　一级分类（数码产品为例）页面

图 14-114　二级分类（消费数码为例）页面

图 14-115　三级分类（录音笔为例）页面

图 14-116　产品（京华录音笔为例）页面

图 14-117　在产品列表页面上选择商家过滤条件

图 14-118　在产品列表页面上选择品牌过滤条件

图 14-119　商家介绍页面

图 14-120　搜索结果页面

图 14-121 “关于我们”页面

14.5.1 涉及全站的调整

1．产品分类稍作调整

18 个一级分类，放在左侧导航中大致为一屏，方便用户浏览。

尽量将各分类下的次级分类及最终产品数均衡。例如，原图书一级分类下有以下二级分类：文学小说、人文社科、经济管理、教育技术、工具书、生活娱乐、外文原版。而图书分类下产品数众多，二级分类偏少，使产品列表翻页过多，不利于收录。建议在图书分类下增加二级分类，如建筑、电脑、计算机、教育、英文、医学等，既能使分类和浏览更准确，也使到达产品页面的点击距离最短。

二级分类下产品多的话，尽量再细分为三级分类。增加分类，看似使某些产品离首页更远了一层，其实是使绝大部分产品离首页的总体距离大大缩短。

2．顶部导航

目前首页顶部中央是 logo 及首页、商家导购（连向商家列表）、分类导购（连向产品分类列表）、导购社区（连向论坛）4 个导航链接，只有 4 个选择，却占据了过大空间。首页第一屏应该迅速展现主题内容。另外，商家导购、导购社区不是网站主体部分，放在这里浪费空间和链接权重，也不利于用户快速寻找产品。

建议改为横贯页面的产品分类导航条，链接指向图书、化妆护肤、服装、数码等一级分类页面。视页面宽度，能放下几个就放几个。最后一个为“所有分类”，链接至目前“分类导购”页面。

顶部导航条设计为下拉菜单，CSS 控制，不可用 JS，用户鼠标放到一级分类名称时，下拉列出最多 10 个二级分类，最后一个是“更多”，链接至本一级分类首页。

这样，搜索引擎蜘蛛和用户都可以迅速沿着链接进入主体，也就是各产品分类。

3．搜索框

目前首页搜索框位置太靠下，不明显，移到顶部导航条下，如其他页面一样。搜索框占用空间也太大，高度应减小。搜索框本身及热门搜索链接可放在同一行，减少空间，也不影响用户使用。

删除“高级购物搜索”链接。流量统计显示，几乎没有人使用高级搜索。

4．H1 文字

所有页面在面包屑导航下、正文内容前加页面标题，用 H1。目前有的页面面包屑导航里的文字做成了 H1，全部取消。

H1 下面加一行简短文字说明，具体文字有的页面需要人工录入（如分类页面），有的是自动生成（如产品页面、搜索页面），下面有具体说明。

5．URL 问题

所有 URL 还是需要静态化。除了各分类页面、产品页面，也包括不容易处理的产品列表页面（三级分类页面）上的按商家、品牌过滤及各种排列方式。

目前有大量相同页面却有不同 URL，造成复制内容。如：

http://www.yicike.com/1169-GSMshouji-/248102-n2680s.htm （为什么 GSM 手机分类后面有个-，而其他分类没有？取消这个连线符，并做 301 转向）

http://www.yicike.com/1169-GSMshouji-/248102-.htm

http://www.yicike.com/1169-GSMshouji-/248102-xxx.htm （xxx 可以是任意字符）

上面三个 URL 都是诺基亚 N2680S 页面，并且网站上出现了前两个 URL，都有收录，浪费资源，挤占了其他页面的收录机会，并造成复制内容。

并且 http://www.yicike.com/1169-GSMshouji-/248102.htm 这个 URL 显示的又与 http://www.yicike.com/1169-GSMshouji-/ 一样。

建议：

无论目录还是 htm 文件，结尾处的连词符“-”一律删除。

http://www.yicike.com/1169-GSMshouji-/ 301 转向至

http://www.yicike.com/1169-GSMshouji/

http://www.yicike.com/1169-GSMshouji-/248102-.htm 301 转向至

http://www.yicike.com/1169-GSMshouji/248102.htm

产品页面 URL 一律只包含产品 ID 编号，删除从产品名称中提取的英文或数字字符。

http://www.yicike.com/1169-GSMshouji-/248102-n2680s.htm 以及

http://www.yicike.com/1169-GSMshouji-/248102-xxx.htm 301 转向至

http://www.yicike.com/1169-GSMshouji/248102.htm

这些 URL 的改动和合并需要做 301 转向。由于 URL 中包含分类编号和产品编号，应该可以根据编号做判断并做转向。

不同页面 URL 格式下面还有相关说明。

6．热门搜索链接

视页面宽度允许，搜索框右侧的（目前是下面，如前述，移至搜索框右侧，同一行）热门搜索增加至 7～10 个，连至静态 URL 的搜索页面。

不同页面使用不同的热门搜索词。如首页是全站热门产品搜索，分类页面是本分类之内的热门搜索，产品页面是本身所在分类的热门搜索。不可全站都是一样的热门搜索，而且最好能轮换，使搜索引擎能抓取更多搜索页面。所以数据库中热门搜索词需要按产品分类划分类别。

热门搜索词可以来源于几个方面：

- 前面关键词研究提到，部分一级分类有很多“地名+产品名”格式的搜索词，如“北京鲜花”，“上海礼品”，“深圳手机”，“北京家电”，“北京商品”，“北京办公用品”等。人工录入数据库。做分类关键词研究时，看到这类不好归入某分类，搜索量又比较大的，都可以做成搜索页面。
- 关键词研究中，搜索次数较少，不能作为分类页面主要目标关键词的，如“促销家电”，人工录入。
- 二级和三级分类页面，因为数量巨大，我没时间做更深入的关键词研究，需要 SEO 人员在 Google 关键词工具里，填写本分类名称，生成相关关键词列表。除了两三个搜索次数最多的词作为本分类页面的主要目标关键词，写入页面 Title，其他搜索次数少的可以录入数据库，作为二级和三级分类页面的热门搜索链接。
- 记录用户实际搜索的词。在不同页面做的搜索记录在相应分类中。搜索词先记录下来，不立即上线，需要人工在后台审核一下，以防搜索词太不相关。
- 产品页面，不同产品可以随机调用不同的（但还是本三级分类下的）热门搜索词，这样更多热门搜索词有机会出现。

搜索页面 URL 和页面内容详见后面。

左侧导航最下面也增加热门搜索。提取本分类搜索次数最多的 20 个热门搜索，做成链接至搜索页面。这里说的左侧导航热门搜索与搜索框下的热门搜索链接处理方法一样，搜索框下用 7～10 个，左侧导航处继续从数据库中提取不同的 20 个搜索词。

7．增加产品资讯新闻板块

新增加资讯、新闻部分，放在/info/或/news/之类的目录下。优惠券及促销资讯纳入这个资讯板块。

从商家网站转载更多优惠券和促销信息，并从商家网站转载、收集更多产品资讯性内容、产品评测。

资讯内容也按产品对应的结构分类。因三级分类数量太大，资讯只分到二级分类即可。面包屑导航及 URL 都按分类明确某个资讯页面所在位置。如某条属于电视机的资讯信息的面包屑导航可以是：

主页 – 资讯 – 家电 – 大家电 – 信息标题

URL 则是：

/info/1725/article123.htm

每个资讯页面按所属分类在页面底部调用本分类热销产品 7～10 个。

8．广告位

网站所有页面按常见广告格式（banner,button,skyscrapper 等）预留广告位。广告管理后台可以控制广告发布到哪些页面，如所有分类页面或全站。

在“关于我们”部分加一个广告信息页面，列出不同级别广告价格。

目前没有广告时先放自己网站的广告，链接到热门分类、促销产品等。

目前除 Google Adsense 等联盟性质的广告外似乎没有商家直接买的广告。在流量达到一定水平后，直接卖广告也是一个很可能的收入来源。建议提前准备。

9．产品属性过滤页面

产品列表页面，通常也就是三级分类页面（有时是二级分类，总之是出现产品列表的页面），按产品属性，也就是各种过滤条件，生成不同产品过滤页面，在左侧列出链接。

首先最明显和简单的是按价格、商家、品牌过滤。目前这三个过滤已有，但有些问题。

以“电视机”分类为例：http://www.yicike.com/1725-dianshiji/，如图 14-122 所示。

图 14-122　电视机分类页面

用户点击商家“168 订购网”后出现如图 14-123 所示页面。

图 14-123　电视机分类商家过滤页面

几个问题：

- 按价格过滤目前是用户填写价格范围，单击“过滤”按钮。这样，页面上不存在用户和蜘蛛可以点击、爬行的链接。建议列出各个价格范围，做成链接至相应价格过滤页面。
- 左侧列出的商家和品牌过滤条件显然有技术问题。商家只列出一个。列出的品牌不是电视机的品牌，似乎所有产品列表页面列出的品牌都是“3M”。
- 用户点击商家过滤条件后，列出的产品似乎不准确，并不限于所选商家。对 SEO 影响倒不大。
- 过滤页面 URL 没有静态化，而且太长、太复杂。

- 用户/蜘蛛选择过滤条件后，页面 Title、面包屑导航、正文顶部说明文字都不能表现出已经选择了过滤条件，页面上唯一的区别是产品数减少了。这无法让用户和蜘蛛辨认内容的区别，更不能突出过滤页面应该针对的关键词。如电视机分类，“168 订购网”商家过滤，本来目标关键词是“168 订购网电视机”，但过滤页面上的 Title、面包屑导航、说明文字都没有“168 订购网电视机”这个关键词。

建议修改如下。

（1）按价格过滤（数字只是举例，下同）

500-1000 元，链接至 http://www.yicike.com/1725-dianshiji/500-1000.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 > 500-1000 元。
- H1 文字：500-1000 元电视机。
- Title：500-1000 元电视机价格、最新报价、评价评测 – 亿赐客比较购物网。

1001-2000 元链接至 http://www.yicike.com/1725-dianshiji/1001-2000.htm

2001-4000 元链接至 http://www.yicike.com/1725-dianshiji/2001-4000.htm

4000 元以上链接至 http://www.yicike.com/1725-dianshiji/4001.htm

（2）按品牌过滤

康佳，链接至 http://www.yicike.com/1725-dianshiji/kongka.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 > 康佳
- H1 文字：康佳电视机
- Title：康佳电视机价格、最新报价、评价评测 – 亿赐客比较购物网

索尼，链接至 http://www.yicike.com/1725-dianshiji/sony.htm

长虹，链接至 http://www.yicike.com/1725-dianshiji/changhong.htm

（3）按商家过滤

京东商城，链接至 http://www.yicike.com/1725-dianshiji/360buy.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 > 京东商城。
- H1 文字：京东商城电视机。
- Title：京东商城电视机价格、最新报价、评价评测 – 亿赐客比较购物网。

1 号店，链接至 http://www.yicike.com/1725-dianshiji/1haodian.htm

世纪电器网，链接至 http://www.yicike.com/1725-dianshiji/51mdq.htm

另外，大多数产品还可以再按某种参数过滤，如电视机还可以按尺寸过滤：

21 英寸以下链接至 http://www.yicike.com/1725-dianshiji/21.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 >21 英寸以下。
- H1 文字：21 英寸以下电视机。
- Title：21 英寸以下电视机价格、最新报价、评价评测 – 亿赐客比较购物网。

21～24 英寸，链接至 http://www.yicike.com/1725-dianshiji/21-24.htm

25～30 英寸，链接至 http://www.yicike.com/1725-dianshiji/25-30.htm

30 英寸以上，链接至 http://www.yicike.com/1725-dianshiji/30.htm

电视机还可以按功能过滤：

等离子，链接至 http://www.yicike.com/1725-dianshiji/plasma.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 > 等离子。
- H1 文字：等离子电视机。
- Title：等离子电视机价格、最新报价、评价评测 – 亿赐客比较购物网。

液晶，链接至 http://www.yicike.com/1725-dianshiji/lcd.htm

普通，链接至 http://www.yicike.com/1725-dianshiji/putong.htm

不同的产品分类，需要不同的参数，不知道是否能自动检测生成，否则需人工定义，可能还需要数据库结构变化，涉及上千个三级分类，工作量不小。将组合产生大量长尾关键词，如上面的“等离子电视机”等。

使情况更为复杂的是，用户可能选择多项属性，如用户点击了电视机 – 索尼 – 京东商城 – 液晶，此时 URL 就需要是：

http://www.yicike.com/1725-dianshiji/sony-360buy-lcd.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 > 索尼 > 京东商城 > 液晶。
- H1 文字：京东商城索尼液晶电视机。
- Title：京东商城索尼液晶电视机价格、最新报价、评价评测 – 亿赐客比较购物网。

URL 中的顺序需要统一和固定，如上面的“品牌—商家—功能”顺序，无论用户是按电视机 – 索尼 – 京东商城 – 液晶顺序来到这个页面，还是按电视机 – 京东商城 – 索尼– 液晶顺序来到这个页面，URL 需要是一样的。所有属性定义优先顺序，URL 按固定顺序生成。

同样，面包屑导航，H1 文字和 Title 也需要定义固定顺序。

在某个过滤条件被选择时，这个属性右侧加一个“取消”链接，点击这个链接将取消选择这个过滤条件。

按属性生成过滤页面的逻辑比较复杂，但很重要。一是对用户有益，直接点击就可以找到自己想要的产品。对 SEO 更重要的是，过滤页面会组合出大量有意义的关键词，如“京东商城电视机”、“索尼液晶电视机”，“2000 元康佳电视机”，“京东商城 21 英寸康佳电视机”等，其中有些搜索量相当大，如“索尼电视机”之类。这种方式组合生成的页面关键词很难用其他方式优化。

10．商家列表及商家介绍页面

商家有关页面分为几类。

（1）商家列表页面

主商家列表页面（旧版网站称为商家导购）http://www.yicike.com/merchant_ranking.htm，如图 14-124 所示。

图 14-124　主商家列表页面

建议新建下级（按产品分类）商家列表页面。

商家列表页面相当于商家网站地图，需链接至各商家介绍页面。人工选出 20～30 个最主要商家，列在主列表页。其他商家按产品分类，列在下一级（对应一级产品分类）列表中。如/merchant/1000-tushu.htm，列出所有网上书店。

主列表页：

- 面包屑导航：首页 > 所有商家。
- H1 文字：网上商城所有商家。
- Title：网上商城所有商家列表评测 – 亿赐客比较购物网。

按分类商家列表页：/merchant/1000-tushu.htm

- 面包屑导航：首页 > 所有商家 > 图书。
- H1 文字：图书网上商城。
- Title：图书网上商城及商家网店列表评测 – 亿赐客比较购物网。

将商家再分类的原因和前面过滤条件页面相同，可以组合生成“图书网上商城”、“家电网上商城”等关键词，用专门页面优化。首先方便用户浏览，现有商家导航页面只能列出很小一部分商家，实际上商家数目远远超过，用户没有简单方法到达所有商家页面。也给搜索引擎蜘蛛预备好更多有明确目标关键词的页面。而且实现并不困难，只要挖掘自身数据库就可以。

目前商家导航页面底部列出了按地点过滤，是个很好的做法，因为如“北京网上商城”这种词也有人搜索。不过目前的实现方法有问题，用户点击地名后，JS 调用了所选地区的商家显示在页面上，URL 却未变化，没有生成新页面，如图 14-125 所示。

图 14-125　商家导航页面上的按地名过滤

改为链接至不同的 URL，如点击北京，链接至 /merchant/beijing.html，在北京页面上列出北京地区或可以发货至北京的商家。

商家按地名过滤页：/merchant/beijing.html

- 面包屑导航：首页 > 所有商家 > 北京。
- H1 文字：北京网上购物商城。
- Title：北京网上购物商城及商家网店列表评测—亿赐客比较购物网。

（2）商家介绍页面

也就是目前 http://www.yicike.com/merchanthome/ejia.htm 这种页面（如图 14-119 所示）。

除商家基本信息外，列出所有这个商家有产品的商家产品分类页面（见下）。如京东

商城在大家电、电脑、数码产品、家具分类都有产品，则列出这些商家产品分类页面。这些分类页面可以列在“产品列表”选项卡下。

- Title：E 家网购物商城网站介绍，优惠券，客服电话，地址 （原因见 14.4.4 节的关键词分析）。
- 面包屑导航：首页 > 商家 >E 家网。
- H1 文字：E 家网网站详细介绍。

（3）商家评论页面

目前商家介绍页面上的“商家评论”选项卡做成链接至独立的商家评论页面，如

http://www.yicike.com/merchanthome/ejia-comments.htm

这个页面允许用户发表评论。

- Title: E 家网怎么样—E 家网评价。
- 面包屑导航： 首页 > 商家 >E 家网评论。
- H1 文字：E 家网用户评价。

在页面正文（用户评论）前的标题或说明文字处，加上“E 家网用户评价”、“您觉得 E 家网怎么样？欢迎提交评论”之类的文字。

（4）商家产品分类页面

一级和二级分类页面（还没有产品列表的页面），左侧导航列出本分类下有产品的商家。如京东商城在家电、电脑、数码产品、家具分类都有产品，则上述分类页面分别列出这些链接：

yicike.com/1009-jiadian/360buy.htm（家电—京东商城）

yicike.com/1007-diannao/360buy.htm 商家一级分类页面（电脑—京东商城）

yicike.com/1008-shumachanpin/360buy.htm 商家一级分类页面（数码—京东商城）

yicike.com/1012-jiaju/360buy.htm 商家一级分类页面（家具—京东商城）

- 面包屑导航：首页 > 家电 > 京东商城。
- H1 文字：京东商城家电。
- Title：京东商城家电价格、最新报价、评价评测 – 亿赐客比较购物网。

商家一级分类页面（家电-京东商城）则列出商家二级分类页面，如 yicike.com/1194-dajiadian/360buy.htm（大家电-京东商城）

在商家二级分类页面列出的是前面讨论过的产品按商家过滤页面如：

http://www.yicike.com/1725-dianshiji/360buy.htm（电视机—京东商城）。

商家分类页面既出现在一级和二级产品分类页面左侧导航，也出现在商家介绍页面的“产品列表”选项卡中。

前面讨论的产品按商家过滤页面出现在三级分类页面（产品列表页面），商家分类页面其实就是产品按商家过滤页面在一级、二级分类的体现。

11．品牌页面

目前没有品牌页面。与商家页面结构类似，给品牌建立一套单独页面。

（1）品牌列表页面

与商家列表页面类似，如：

www.yicike.com/brands.htm 品牌主列表页面

www.yicike.com/brands/1000-tushu.htm 按分类列表页面

（2）品牌介绍页面

使用类似 http://www.yicike.com/brand/hp.htm 这种 URL。

- Title：惠普报价、专卖店价格、产品信息。
- 面包屑导航：首页 > 品牌 > 惠普。
- H1 文字：惠普报价及产品信息。

页面内容也与商家页面类似，除品牌基本信息外，列出这个品牌有产品的品牌产品分类页面（见下）。如惠普在家电、电脑、数码产品等分类都有产品，则列出这些品牌产品分类页面。

（3）品牌评论页面

在品牌主页面上加链接到品牌评论页面，如：

http://www.yicike.com/brand/hp-comments.htm

这个页面允许用户发表评论。

- Title：惠普产品评测、用户评论。
- 面包屑导航：首页 > 品牌 > 惠普评论。
- H1 文字：惠普用户评价。

（4）品牌产品分类页面

与商家产品分类页面类似。一级和二级分类页面（还没有产品列表的页面），左侧导航列出本分类下有产品的品牌。如家电分类下：美的家电，西门子家电，三星家电……

美的家电为例：yicike.com/1009-jiadian/meidi.htm

- 面包屑导航：首页 > 家电 > 美的。
- H1 文字：美的家电。
- Title：美的家电价格、最新报价、评价评测 – 亿赐客比较购物网。

电脑分类下：联想电脑，三星电脑，惠普电脑。

化妆护肤：dhc 化妆品，资生堂化妆品，雅芳化妆品，迪奥化妆品。

数码：索尼数码，佳能数码，三星数码，松下数码。

食品：统一食品，百事食品，雨润食品，达利食品，光明食品。

手机通讯：诺基亚手机，三星手机，索爱手机，联想手机，多普达手机。

品牌产品分类页面则列出产品按品牌过滤链接，如康佳电视机 http://www.yicike.com/ 1725-dianshiji/kongka.htm

（5）CSS

除了外置的 CSS 文件，页面 HTML 中还是有大量 CSS 代码，删除或集中到外置 CSS 文件。直观从 HTML 代码看，div 类的代码非常多，尽量删除，可能可以使页面缩小 50%，也许更多，减少干扰，也使页面访问速度更快。

JS 代码，如果是多个页面常用的，也尽量放在外部文件。不能用外部文件时也尽量放在代码底部。

12．网站 Logo

所有页面左上角 Logo 做成连至首页的链接。ALT 文字："网上购物网站排名问网购专家-亿赐客比较购物搜索网"。图片 ALT 文字相当于链接锚文字，为避免过度优化，与首页 Title 稍作区别。

13．注册和免费订阅

我尝试注册，但一直显示"验证码错误"（我肯定验证码输入是正确的），所以不知道注册用户账号里的功能有什么。免费订阅注册后也没有收到什么邮件。

这两个功能建议合并。如果需要做邮件营销，向注册用户发邮件就可以，不必做成两个数据库。在吸引浏览者注册方面还需要改进，目前看不出注册有什么好处。

吸引用户注册是非常重要的，比较购物网站的致命缺陷之一就是没有真正属于自己的用户。

14．页脚

亿赐客网站页脚如图 14-126 所示。

首页页脚增加链接至使用条款 terms.htm，隐私权政策 privacy.htm。其他页面页脚不要这两个链接，以免浪费权重。

图 14-126　亿赐客网站页脚

以下指全站页脚。

“比较指南”部分全部删除。可以放上一个“帮助”或“用户指南”之类的链接，连至常见问题列表。但没有必要在全部页面上列出多个问题，浪费权重。用户有不明白的地方，有一个“帮助”链接已经足够。

“网站导航”部分：

- “导购社区”链接移至页头右侧，“使用帮助”前。
- “分类导购”文字改为“产品分类”。“商家比较”改为“所有商家”。这样意义更明确。
- 增加“网购资讯”，链接至资讯部分首页。
- 增加“所有品牌”，链接至品牌列表页面。
- 广告服务、人才招聘链接删除，在“关于我们”页面加上这两个页面链接，没有必要出现在所有页面。

版权声明部分，加中文“亿赐客比较购物搜索”并链接至首页。

15．网站目标及导出链接跟踪

目前 GA 中设置的转化目标是实现一次站内搜索。这并不是一个适合亿赐客的网站目标。目前亿赐客的主要赢利模式是联署计划（网站联盟），通常这种网站目标应定为用户点击导向商家网站的联署链接。亿赐客网站本身的目的就是吸引、推动用户点击联署

链接，点击越多，获得佣金的机会越大。用户到达商家网站后怎样转化，则是商家网站需要研究的问题。

为了记录、研究、优化网站转化，所有导向商家的链接都需要做点击跟踪。GA 中记录点击（到其他网站）并把点击设置为转化的说明：

http://www.google.com/support/googleanalytics/bin/answer.py?hl=en&answer=55527

http://www.google.com/support/analytics/bin/answer.py?hl=en&answer=72712

今后网站扩展其他赢利模式，大致上也应该是以点击到其他网站次数为基础的，如直接卖给广告商的显示广告、按点击收费等，有了 GA 记录的点击数字，就可以提供潜在广告数据，说服广告商。

16．分类页面友情链接

除首页底部友情链接外，所有分类页面也在底部留出友情链接位置，并在后台管理。除首页外，分类页面也可以交换链接。从三级分类页面开始交换深层链接，而不是一级分类。当然不是绝对的，只是将大部分时间和人力首先分配到三级分类页面。

交换链接时，锚文字交替使用。比如首页就可以交替使用：亿赐客、亿赐客购物搜索、亿赐客网上购物、购物网站比价、亿赐客网购比价、亿赐客比较购物、购物网站评测，等等。分类页面也如此，锚文字不要仅用分类名称，多一些变化形式，交替使用。

17．网站名称

使用“亿枝客”有什么特殊考虑？为什么不是与域名拼音相符的“亿次客”？最好不要使用户有任何混淆的机会，好不容易记住网站名称，打入拼音，却找不到网站。

14.5.2 首页修改

除了前面讨论的涉及全站和网站结构方面的改动外，还建议进行如下修改。参考图 14-127，简陋的手绘首页示意图。

Title：网上购物网站谁最好？问网购专家—亿赐客比较购物搜索网。

- 融合 3 个最主要关键词：网上购物、购物网站、网购，又自然组合出“网上购物网站”（参考 14.4.1 节“首页关键词分析”）。
- 以问句吸引用户点击。
- 包含品牌名称。
- 结尾处出现“比较购物”、“购物搜索”，为在行业内建立品牌做准备。

Meta Description：网上购物网站哪家最好？哪里价格最便宜？网购专家—亿枝客比

较购物搜索网帮你省钱省时间。千家网上商城，数百万商品最新报价，促销优惠券，用户评论，购买窍门，尽在亿赐客。

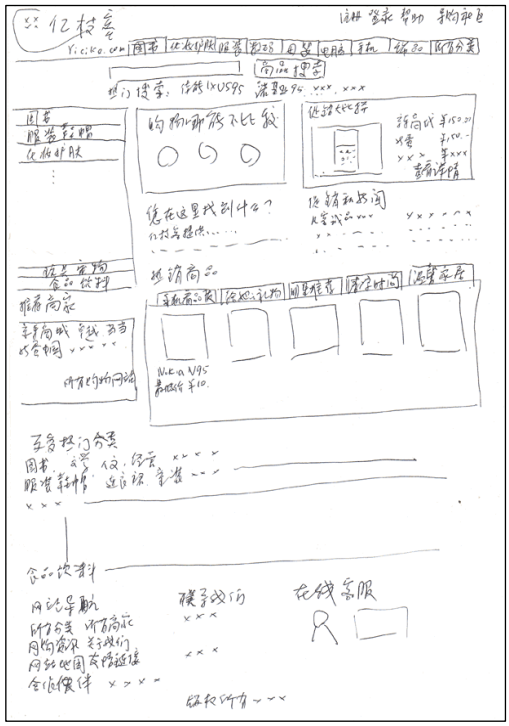


图 14-127　首页修改示意图

- 出现目标关键词。主要作用是吸引点击。

页面顶部 logo 位置改到左上角，与其他页面一致。左上角口号“网上购物，先比再购”删除。Logo 右侧放文字“网上购物网站谁最好？问网购专家亿赐客！”，使用 H1。

顶部导航见全站顶部导航说明。顶部导航下为搜索框。

左侧加 CSS 导航，列出所有一级分类页面及 10 个二级分类页面。

左侧导航下方，推荐商家链接 20 个，连至相应商家介绍页面。最后是“所有购物网站”，连至商家主列表页面。不宜列出过多，首页权重应该尽量导向网站主体架构：分类页面和产品。

左侧的“产品排行榜”全部删除。排行榜中的链接其实大部分是商家介绍页面，已有推荐商家代替。

将目前搜索框下的“推荐商品分类”和“推荐购物网站”两个选项卡及其内容全部删除。

卖点框（购物哪能不比较）和促销大比拼占用太多空间，尺寸缩小，右移。促销大

比拼部分，“查看详情”只留一个，没有必要以“查看详情”这样的链接分散锚文字相关性。商家图片链接直接通过联署链接导出至商家网站（而不是本网站上的产品详情页面），尽快完成网站目标。

卖点框下面是网站说明文字，作用相当于目前底部的“我们在做什么”部分。标题改为“您在这里能找到什么？”，文字尽量融合：网上购物、网购、购物网站、购物网、团购、代购、淘宝、店铺、商品、价格、旺旺、服务、图片、购买、钻石、朋友。目前的分类名称罗列不必要。

参考文字：

简单说，最快速地找到最低价格——网上购物价值最大化。上千购物网站，数百万商品，从淘宝店铺到 B2C 购物网，从个人网购到团购代购。数据及时更新。全面、公正、透明的网上购买体验，从亿赐客开始。

说明文字右侧是“促销和购物新闻”链接 10 条左右，连至资讯新闻内容。

热卖商品以选项卡形式实现，大幅节省页面空间。“最低商家”语义不清，改为“最低价商家”。商家名称直接链接到商家网站，而不是商家介绍页面，尽快完成网站目标。不要放促销信息，热卖商品处的促销信息链接过多，浪费宝贵的首页链接权重。

更多热门分类，类似目前首页“推荐商品分类”部分，但应该更有条理。每个一级分类一行，一级分类名称列在最左侧，右面列出左导航 CSS 菜单中没有列出的二级和三级分类名称。二级和三级分类用不同颜色字体，以示区别。

视页面美观与否，“更多热门分类”也可以分为两列，这样 18 个一级分类，只需要 9 行，不至于失衡。

14.5.3 一级分类页面

一级分类页面指这类页面：http://www.yicike.com/1008-shumachanpin/，参考图 14-113。

目前一级和二级分类页面不知为什么侧导航放在了右侧，看不出有什么特殊益处。建议全站统一，全部放左侧，或全部放右侧。

H1 文字紧接着面包屑导航下面： xxx 价格/报价 （xxx 为分类名称，下同），如“数码产品价格/报价”。

H1 下面一行说明文字需要根据关键词研究人工撰写，自然融合搜索次数最多的 6～8 个相关关键词。目前的说明文字：

亿枝客数码产品频道：提供数码产品价格，行情，数码产品报价，最新数码产品导购，数码产品图片，新闻，数码产品评论，产品参数，帮助你挑选出最满意的数码产品。

稍有关键词堆积之嫌，并且没有出现其他主要的相关关键词。应根据关键词研究重写，如 Google 关键词工具给出数据，如图 14-128 所示。

可考虑说明文字写为：

数码产品报价，包括数码摄像机、照相机、MP3 等商品，数百家数码商城最新价格比较，及产品配件，图片，用户评论，帮你找到最低价的数码产品。

请注意写法的不同：

- 不仅要包含“数码产品”，也要包含数码摄像机、数码商城、配件等搜索次数多的相关关键词。“数码摄像机”这种词不是要用这个页面优化，它应该有自己相应的二级分类页面，放在这里是为了支撑“数码产品”的语义相关性。
- 不必多次重复“数码产品”，出现两三次就足够了。
- 比 Title 包含更多关键词，Title 中出现的关键词必须出现在这段说明文字中。
- 可以加一些无关的词语（如“数百家”，“帮你找到”这类的词），使句子自然。

再举一例，“食品饮料”分类如图 14-129 所示。

	A	B	C	D
1	关键字	广告竞户竞争程度	本地搜索量：6月	全球每月搜索量
2	数码摄像机	0.93	673000	450000
3	数码产品	0.93	201000	201000
4	数码照相机	0.86	165000	165000
5	数码商城	0.66	27100	27100
6	数码产品配件	0.46	-1	3600
7	数码相机产品	0.53	6600	1900
8	数码产品报价	0.53	1900	1900
9	数码产品商城	0.46	-1	880
10	数码产品销售	0.4	880	880
11	数码产品代理	0.6	590	720
12	代理数码产品	0.46	590	720
13	索尼数码产品	0.53	1300	720
14	数码产品 深圳	0	-1	590
15	深圳数码产品	0.4	-1	590
16	mp3数码产品	0.4	880	390
17	数码产品mp3	0.46	880	390
18	数码产品价格	0.53	-1	390
19	北京数码产品	0.4	480	390

图 14-128　数码产品关键词

	A	B	C	D
1	关键字	广告竞户竞争程度	本地搜索量：6月	全球每月搜索量
2	食品	1	4090000	3350000
3	食物	0.8	1000000	823000
4	饮料	0.86	673000	450000
5	食品饮料	0.73	90500	90500
6	食品网	0.6	90500	90500
7	食品机械	0.86	74000	74000
8	绿色食品	0.8	60500	60500
9	食品批发	0.6	90500	49500
10	休闲食品	0.8	60500	40500
11	进口食品	0.66	33100	40500
12	饮料机械	0.8	14800	33100
13	日本食品	0.66	9900	33100
14	餐饮食品	0.4	-1	33100
15	广东食品	0.46	40500	27100
16	小食品	0.73	22200	27100
17	天津食品	0.46	40500	22200
18	深圳食品	0.33	22200	22200
19	广州食品	0.26	-1	14800

图 14-129　食品饮料关键词

去除不太相关的“食品机械”等，说明文字可写为：

食品饮料网购最低价格，数百家食品网站最新报价，绿色休闲食品批发零售，汇集全国各地如广东，深圳，天津，台湾及进口食品。

不同分类有不同关键词，说明文字写法也不同，不能用统一格式自动生成。在完成人工撰写之前，暂时先用格式套用：

亿赐客 xxx 频道:提供 xxx 价格，行情，最新报价，导购资讯，图片，新闻，评论，产品参数，帮助你挑选出最满意的 xxx。

左侧栏：

- CSS 控制右拉菜单：列出所有二级分类，及每个二级分类下 5～7 个三级分类。
- 推荐商家：列出本分类有产品的商家 20 个，文字链接，链接到相应的商家产品分类页面（不是商家介绍页面）。最后是“所有商家”链接，链接至本分类所有商家列表。
- 推荐品牌：列出本分类有产品的品牌 20 个，文字链接，链接到品牌产品分类页面。最后是“所有品牌”链接，链接至本分类所有品牌列表。
- 分类资讯：列出新增加的与本分类相关的资讯链接 5 条。相当于目前右侧导航底部的“最新优惠券”、“最新促销资讯”，但不用 10 条这么多，以免过多分散权重。只列出本分类下的资讯，而不是目前这样全站列出的都是相同的 10 条资讯，不然这 10 个资讯页面获得了全站链接及没有必要的高权重。
- 热门搜索：列出本分类相关的搜索链接 20 个。

热卖产品：目前页面最上部产品。照片、产品名称、“详情”链接到产品页面，但注意 URL 需统一，目前图片和产品名称的链接不是一个版本的 URL。“最低价格”下面的“在 xxx 去购买”用联署链接到商家网站，同目前“价格比较”下商家 logo 链接一样，而不是链接到商家介绍页面，文字改为“到 xxx 去看看”。不要直接放联盟链接，和产品页面一样，打开新窗口，使用脚本转向，并用 Google Analytics 跟踪点击次数，记为一次转化。

正文部分 xxx 分类列表：热卖产品下面列出所有下级分类，包括二级和三级。

最新产品 （原推荐产品）：15～20 个产品，只需列出图片（链接至产品页面）、产品名称（链接至产品页面）、最低价格及商家（联署链接至商家网站）。“比较价格”和“降价通知”没有必要。

H2 文字：xxx 产品分类，推荐商家，推荐品牌，分类资讯，热门搜索，热卖产品，xxx 分类列表，xxx 最新产品，这些小标题使用 H2。

Title：针对分类名称做关键词研究，然后人工撰写 Title。下面列出我根据关键词搜索次数写的 Title，供参考。

- 图书分类 Title：图书批发购买，网上书店特价图书音像价格—亿赐客比较购物。
- 娱乐分类 Title：最新游戏，最新音乐，电影娱乐产品价格—亿赐客比较购物。
- 服装鞋帽分类 Title：品牌服装，韩国服装，时尚流行服装价格—亿赐客比较购物。
- 珠宝饰品分类 Title：珠宝饰品，休闲流行时尚服饰，韩国服饰价格—亿赐客比较购物。
- 礼品鲜花分类 Title：鲜花速递，网上订花，礼物工艺品，礼品网价格—亿赐客比较购物。
- 手机通讯分类 Title：智能手机报价，最新手机配件价格—亿赐客比较购物。

- 电脑分类 Title：笔记本电脑报价，特价手提电脑价格，电脑配件—亿赐客比较购物。
- 数码产品分类 Title：数码产品报价，数码商城产品价格—亿赐客比较购物。
- 家电分类 Title：家用电器，品牌家电报价，小家电价格—亿赐客比较购物。
- 化妆护肤分类 Title：化妆品价格，韩国日本化妆品，品牌化妆品护肤品—亿赐客比较购物。
- 汽车用品分类 Title：汽车用品报价，汽车美容装饰用品，配件及维修—亿赐客比较购物。
- 家居园艺分类 Title：家居用品，园艺用品，时尚智能家居装饰饰品—亿赐客比较购物。
- 母婴用品分类 Title：母婴用品，儿童婴儿用品，宝宝用品报价—亿赐客比较购物。
- 运动户外分类 Title：户外运动服装，运动器材，运动休闲服饰—亿赐客比较购物。
- 健康医药分类 Title：健康医药产品，医疗器械，医疗保健产品报价—亿赐客比较购物。
- 办公用品分类 Title：公司办公用品批发，办公家具文具设备报价—亿赐客比较购物。
- 食品饮料分类 Title：食品饮料，绿色休闲食品，进口食品价格—亿赐客比较购物。
- 玩具宠物分类 Title：模型玩具，毛绒玩具，宠物用品价格—亿赐客比较购物。

人工写 Title 注意几点：

- 整个 Title 最多 30 个汉字。
- 选 2～3 个搜索次数最多的词，搜索次数最多的排在最前面。
- 不必刻意重复分类名称。
- 句子需尽量通顺。
- 如果能把关键词连起来，组合在一起，可以包括更多关键词。比如化妆护肤分类，搜索次数多的关键词包括：化妆品、日本化妆品、化妆品价格、品牌化妆品、韩国化妆品，整合为“化妆品价格，韩国日本化妆品，品牌化妆品护肤品—亿赐客比较购物”，3 个短语包括了 5 个关键词，而且读起来也还通顺。

14.5.4 二级分类页面

二级分类页面指这类页面：http://www.yicike.com/1186-xiaofeishuma/ ，参考图 14-114。

二级分类页面与一级分类几乎相同。

H1，紧接着面包屑导航下面：xxx 价格/报价。

H1 下面一行说明文字根据关键词研究人工撰写，自然融合搜索次数最多的 6～8 个相关关键词。写法参考一级分类页面部分举例和说明。这部分很花费时间，可以慢慢做，

非紧急。

左侧导航中的 xxx 产品分类列出 15～20 个三级分类。

正文中 xxx 分类列表列出所有三级分类。

Title：二级分类不算很多，可根据关键词研究人工撰写 Title。慢慢做。

14.5.5 三级分类页面（产品列表页面）

三级分类页面（产品列表页面）指这类页面：http://www.yicike.com/1360-luyinbi/ ，参考图 14-115。

1．左侧栏

- “谁说比较浪费时间”部分：删除，内容没有必要。
- 价格范围：不要使用表格方式，改为列表和链接。见前面产品按属性过滤部分说明。
- 产品功能过滤：不同产品有不同功能或参数过滤条件，如录音笔可以有录音时长、信噪比等。见前面产品按属性过滤部分说明。
- 推荐商家：列出本分类有产品的商家 20 个，文字链接，链接到按商家过滤页面。见前面产品按属性过滤部分说明。
- 推荐品牌：列出本分类有产品的品牌 20 个，文字链接，链接到按品牌过滤页面。见前面产品按属性过滤部分说明。
- 热门搜索：列出本分类的搜索链接 20 个。

2．翻页链接

产品较多的分类会有数百个翻页，按目前 1～10 这种标准翻页链接格式，要访问或爬行到第几百页，得点击几十次。这对搜索引擎蜘蛛来说是不可能的。这是可能造成收录问题的原因之一，被推到后面翻页的产品没有机会被爬行，如图 14-130 所示。

图 14-130　产品列表页面上的翻页链接

建议顶部及底部翻页都分为两排。在目前 1，2，3，……，10 下，加一行：

11，21，31，41，51，61，71，81，91，101

视排版和页面宽度，两行翻页链接都可以不止 10 个。这样，两行翻页有不同步长，将大大减少到达产品页面所需点击数，缩短产品页面距首页的距离。

翻页链接均需静态化，如第一个分类页面是：

http://www.yicike.com/1594-doujiangji/

则第二页是：

http://www.yicike.com/1594-doujiangji-p2/

3．各种显示方式

各种显示方式如图 14-131 所示。

图 14-131　产品列表页面的各种显示方式

（1）按网格方式显示：默认方式。显示按网格方式排序页面时，此格式高亮并设置为非链接（已经在此显示方式，不可点），“按列表方式排”则为可点击的 JS 链接，点击后页面变为按列表方式排，这里的 JS 链接阻止搜索引擎爬行和抓取。

（2）按列表方式排：显示按列表方式排序页面时高亮并设置为非链接，同时“按网格方式排”设置为正常（非 JS）链接。

（3）按评价排：默认方式。显示按评价排序页面本身时，此格式高亮并设置为非链接，按价格、按名称排序设置为 JS 链接，可点击访问，但阻止蜘蛛爬行。

（4）按价格排、按名称排：同上处理。

也就是说，只有按网格方式+评价排的页面（包括其翻页）才被收录，其他所有排列方式都用 JS 链接，阻止搜索引擎爬行。为保险，其他排列方式页面加 noindex，follow 标签禁止搜索引擎收录和索引。

各种显示方式、排列方式页面本身是没有排名意义的（有一个分类页面就足够了），被收录唯一的好处是给产品页面提供更多爬行入口。建议不允许收录除网格+评价排序的列表页面，原因有两方面：

- 目前亿赐客网站 PR 和权重还不太高，能被收录的页面总数可能在几百万页之内。如果允许各种显示方式、排列方式页面都被收录，可能减少其他更应该被收录的页面（如最终产品页面）被收录的机会。
- 只要网格+评价排序列表页面翻页结构解决好，已经可以为产品页面提供爬行入口。

当然，这是针对目前亿赐客网站情况的策略。如果以后网站 PR 值达到高端 6 或 7，能带动的总页面数上升到几千万，也不妨允许收录各种显示方式、排列方式页面，提供更多产品页面入口，进一步提高收录率。

目前默认显示是按价格排，建议改为按评价排序，因为按价格排，第一页有时是价格低但不太相关的产品，比如摄像机分类，排在第一页的是插头、擦布、录像带、玩具摄像机之类的，而不是用户期待看到的摄像机。默认按评价排序还可以在一定程度上人工调整排在第一页的产品，只要人工加几个评价。

图 14-132　产品信息显示

4．显示产品信息

产品信息显示如图 14-132 所示。

- “最低价格：xxx 元”处文字，字号放大，甚至可以用红色，视觉上突出，吸引点击。
- “在 xxx”商家文字，通过联署链接连至价格最低商家网站，而不是目前的商家介绍页面。注意前面提到的网站目标设置和点击跟踪。
- “去购买”按钮：通过联署链接连至价格最低商家网站。
- 产品图片加 ALT 文字，与产品名同。图片通过联署链接连至价格最低商家网站。
- 产品名称不要用 H3。
- 产品名称下面（或后面）加文字，“共 xx 个商家报价”，链接至产品页面。

以上 6 点也适用于搜索页面等处的产品信息显示。

5．面包屑导航

面包屑导航处本分类名称不要用 H1，视实际视觉效果，可考虑改为黑体。

6．H2 文字

左侧导航的“产品参数过滤”、“价格范围”、“推荐商家”、“推荐品牌”、“xxx 相关搜索”等用 H2。

7．H1 文字

紧接面包屑导航下面加 H1 文字：

“分类名称”+价格，最新报价

如录音笔分类就是：

录音笔价格，最新报价

8．说明文字

H1 文字下面或右侧加本分类说明文字。与一级、二级分类一样，根据关键词研究撰写说明文字。

如“豆浆机”分类：

各种家用、全自动、多功能、大型、小型豆浆机价格、报价，配件、维修、二手求购信息

再举一例，“婴儿推车”分类：

婴儿车价格/报价，各种婴儿推车、婴儿手推车、双胞胎婴儿车网上购买信息

9．Title

从关键词研究可以看到，“豆浆机”这个词相关的搜索次数最多的是：九阳豆浆机，豆浆机，全自动豆浆机，豆浆机配件，家用豆浆机，豆浆机价格，大型豆浆机，多功能豆浆机，豆浆机维修，小型豆浆机等。

Title 可写为：

豆浆机价格/报价，全自动、家用、多功能豆浆机产品信息

除了“xxxx 价格/报价”外，再选择搜索次数最多的 2～4 个词。后面不必加“亿赐客比较购物搜索网”。

三级分类有上千个，需要人工做关键词研究并写 Title 和说明文字，熟练后也大概至少需要 5～10 分钟做一个分类，工作量大。不必求快，慢慢做。

人工处理前，暂时先自动按分类名称生成 Title 和说明文字：

- Title：xxxx 价格/报价/评测—亿赐客帮你网上购买最低价的 xxxx
- 说明文字：亿赐客 xxx 频道：提供 xxx 价格，行情，最新报价，导购资讯，图片，新闻，评论，产品参数，帮助你挑选出最满意的 xxx。

翻页后，从第二个页面开始，Title，H1 文字需与第一个页面有所区别。如第二个页面：

http://www.yicike.com/1594-doujiangji-p2/

- Title：第二页 - 豆浆机价格/报价，全自动、家用、多功能豆浆机产品信息
- H1：豆浆机价格，最新报价—第二页

同样，按表格显示，按评价、名称排列的第二页起，虽然使用 JS 链接使搜索引擎不能爬行，但 Title，H1 文字也需与第一个页面有所区别，利于用户分辨自己所在位置。

14.5.6 产品页面

参考图 14-116。

面包屑导航里的产品名称可考虑改为黑体，如果视觉上不太突兀难看的话。

正文最前面的蓝色黑体产品名称放入 H1 中。

产品图片 alt 文字与产品名称同。

价格比较、商家比较、详细资料、查看评论 4 个选项卡位置不变，文字放入 H2 。

价格比较—商家形象下的 Logo，及商家比较—评级里的商家名称，链接到商家网站（通过联署链接及跟踪），不要连到商家介绍页面。红色促销标志也同样。网站的目标是尽快让用户点击联署链接，商家介绍页面没有必要被链接这么多次。

按钮文字“去购买”和“去看看”、“查看详情”等做实验看哪个转化率高。不必用工具，每个实验组放 24 小时就可以看结果。

正文中间的产品品牌、分类、产品人气，分类的所有链接似乎都是连到无效页面，显示 ID NOT Found，修改链接至正确的分类页面。品牌名称链接至品牌分类页面。

如图 14-133 所示，举报及剪贴图标无法辨认，不点击根本不知道是什么。书签图标混淆，不宜使用 RSS 图标。建议使用文字图标。

图 14-133　无法辨认的图标

Title：（xxx 为产品名称）

xxx 价格/报价，xxx 怎么样？好吗？产品评价，图片—亿赐客比较购物搜索网

写法原因见 14.4.5 节“产品页面关键词研究”。

富摘要（rich snippets）：rich snippets 是 Google 前不久推出的一种标签格式，在 HTML 代码中加入标签，Google 搜索结果就将显示标签里的格式化内容。比如论坛帖子就可以显示作者、回复数、最新回复日期等。

产品页面加入 rich snippets 标签，使显示结果更突出，提高点击率。产品页面上的这几个信息适合加入 rich snippets：

- 最低价格。
- 品牌。
- 评论数或产品人气。
- 报价商家数量。
- 技术上可能的话，某些技术参数，如手机待机时间、内存等。但不同产品势必选取不同参数，不知技术上实现难度怎样，难度太大的话，就不必考虑了，前面 4 个就够了。

代码写法参考：

http://googlewebmastercentral.blogspot.com/2009/05/introducing-rich-snippets.html

http://www.google.com/webmasters/tools/richsnippets

14.5.7 产品按属性过滤页面

前面全站修改部分对产品按属性过滤页面已经说明和举例。这里只简单总结。

以“电视机”分类为例：

http://www.yicike.com/1725-dianshiji/

按价格过滤：

500-1000 元，URL 为 http://www.yicike.com/1725-dianshiji/500-1000.htm

- 面包屑导航：首页 > 家电 > 大家电 > 电视机 > 500-1000 元。
- H1 文字：500-1000 元电视机。
- Title：500-1000 元电视机价格、最新报价、评价评测 - 亿赐客比较购物网。

面包屑导航、H1 文字、Title 都需要包括分类名称（电视机）及属性（500-1000 元），由程序自动生成。

这也适用于商家分类页面、品牌分类页面。

14.5.8 搜索页面

搜索页面 URL 也需要静态化。统一采取 www.yicike.com/search-xxx/ 格式。其中 xxx 最好是符合这样条件的英文字符串：

- 不同搜索词，字符串也不同。
- 相同搜索词，字符串也相同（所以 URL 也就相同）。

简单说，每个搜索词对应唯一的 URL，使每一个搜索页面成为一个独立的可被收录的静态页面。这对英文网站来说很简单，xxx 部分替换为搜索词就可以了。中文搜索词则不同。如果技术上实现太困难，也可以直接使用搜索词，中文搜索词也原样用在 URL 中。

不同产品排列方式（网格、列表，价格、评价、名称）的处理与三级分类页面相同，包括使用 JS 链接阻止爬行，Title、说明文字的区分（即使已经阻止爬行）。

翻页处理与三级分类页面不同。所有翻页链接不用正常链接，而是用 JS 脚本实现，与按价格、名称排序的链接相同。也就是说，搜索页面只有第一个页面允许收录，其他

（第二页，第三页，按列表显示，按价格、名称排序）都不让搜索引擎爬行，因为这些页面上的产品都已经在三级分类页面上出现和爬行了，第一个页面用于关键词排名已经足够，其他页面没必要爬行和收录。

重要：与此类似，凡是出现产品列表的地方（因此出现几十、上百个翻页），如商家分类和过滤页面、品牌分类和过滤页面、按参数过滤页面等，都用相同方法处理翻页，使商家分类/过滤页面、品牌分类/过滤页面、参数过滤页面等都是只收录第一个页面，从第二页开始都使用 JS 脚本甚至 AJAX 阻止爬行。结果是，所有产品页面只有一个收录入口，也就是三级分类页面。

搜索页面其他地方大致与三级分类页面相同。

面包屑导航处搜索词不要用 H1。

左侧导航的所有小标题（价格范围、选择分类、选择商家、选择品牌等）用 H2 。

紧接面包屑导航下面加 H1 文字：

“搜索词”+价格，最新报价

如：

冲锋衣价格，最新报价

H1 下面或右侧说明文字：

网上购买 xxxx 最低价格，购物网站最新报价，产品信息，图片，评价评测。

如：

网上购买冲锋衣最低价格，购物网站最新报价，产品信息，图片，评价评测。

页面 Title：

热门搜索：xxxx 价格/报价/评价 – 亿赐客帮你网上购买最低价的 xxxx

如：

热门搜索：冲锋衣价格/报价/评价 – 亿赐客帮你网上购买最低价的冲锋衣

14.6 执行、效果及后续

2009 年 10 月份提交优化建议后，亿赐客团队当月完成了 URL 静态化，收录很快开始增长。2010 年初观察一段效果并考虑一些技术问题后，对 URL 格式又做了两次大规

模变动。例如，原计划产品页面 URL 为：

http://www.yicike.com/1169-GSMshouji/248102.htm

但考虑到有些时候产品可能需要移动至另一个分类（如人工检查、调整产品分类准确性时），上述格式势必产生 URL 变化，因此去掉分类目录，只保留产品编号：

http://www.yicike.com/248102.htm

产品分类编号也做了重新调整。前面优化报告中提到的一级分类：

http://www.yicike.com/1008-shumachanpin/

现在 URL 为：

http://www.yicike.com/100001008-shumachanpin/

因此从编号就可以知道是哪一级分类。

2010 年上半年，由于 URL 系统的变动，收录波动较大。即使做了 301 转向，新旧 URL 很长一段时间同时存在于搜索引擎数据库中，尤其是百度，对 301 转向反应很慢。但从总体上看，做了 URL 静态化和网站链接结构的修改，收录明显快速增加。

2010 年 1 月 1 号，第一个较完整优化版本上线。2 月底，网站流量上升到每天 5 万个 IP 以上。对比网站优化前的流量，这就是本书副标题“60 天，网站流量提高 20 倍”的由来，如图 14-134 所示。

图 14-134　亿赐客网站 2 月份流量

当然，1 月 1 日上线的版本还有错误和不完整，后来的执行过程中也遇到各种问题，计划也会相应变化。例如：

- 原建议书中左侧导航底部的热门搜索、推荐商家、推荐品牌等链接，在一级、二级分类页面上都移到了右侧正文最下面。

- 产品列表页面上的按价格过滤，做成列表及链接后效果不好。不同产品种类有不同的价格区间，价格过滤条件就必须不一样，比如电视机和图书类别的价格过滤条件相差很大。以程序计算最合适的价格过滤条件遇到一些问题，有的分类列出的价格不太靠谱。比如某个商家某本书价格错写为 50 万，就可能使整个图书类别最低价格区间变成 0～1000 元，使整个价格过滤系统失去意义。因此最后又取消了按价格过滤链接，恢复为原来的用户在表格中输入价格。
- 产品按参数或功能过滤，由于比较复杂，工作量大，没有实施。
- 网站链接结构修改过程中，由于一个大家都没注意到的技术细节，使网站实际上生成了上亿页面，当然，都是复制内容。我查看网站日志时，发现搜索引擎蜘蛛在爬行某类 URL 时似乎陷入无限循环，根本没时间爬行其他 URL 了，才发现这个技术问题。
- 品牌相关页面，由于工作量比较大，时间又紧，没有实施，以搜索页面代替。
- 2010 年 3 月 17～19 日，百度有一次算法的改变，使几乎所有比较购物类网站（也包括其他类网站）受到重挫，流量大幅下降，包括亿赐客。5 月初，Google 有一次被称为“May Day”的算法更新，对长尾词影响巨大，亿赐客也受到波及。4～6 月份，我们对亿赐客及新算法又做了分析和调整。7 月份后调整近乎停止，等待搜索引擎重新收录和计算。2010 年 8 月底流量又重新恢复和上升中。

执行、观察效果、发现问题、再次修改，循环往复，再加上竞争对手及搜索引擎算法的不停变化，所有网站的优化都处于不断调整、变动的过程中，没有终止。

SEO 术语

301 转向（301 Redirect）

也称为 301 重定向、301 跳转。

301 转向是用户浏览器或搜索引擎蜘蛛向网站服务器发出访问网址 A 的请求时，服务器返回的 HTTP 数据头信息状态码的一种，表示所请求的网址 A 永久性地转移到另一个网址 B。

301 转向是搜索引擎友好的转向，网址 A 的权重和 PR 值将会被转移到网址 B。所以当网站上有网址变动时，建议使用 301 转向。

302 转向（302 Redirect）

也称为 302 重定向、302 跳转。

302 转向是用户浏览器或搜索引擎蜘蛛向网站服务器发出访问网址 A 的请求时，服务器返回的头信息中状态码的一种，表示所请求的网址 A 暂时性转移到另一个网址 B。除非网页 A 真的是短时间转移到网页 B，可以使用 302 转向，否则不建议使用，302 转向对搜索引擎不太友好。

Alexa 排名

网上最常用、历史最悠久的网站流量排名工具，由亚马逊书店所拥有。在 Alexa 网站输入域名后，Alexa 会以曲线形式显示网站流量趋势以及按流量计算出的世界排名。也可以输入多个域名，Alexa 会把几个域名的流量曲线显示在一起以方便比较。

白帽 SEO（Whitehat SEO）

白帽 SEO 指的是合理合法，有利于用户体验，符合搜索引擎质量规范的 SEO 手法。通常使用白帽 SEO 手法排名会比较持久稳定。

标签（Tag）

标签指的是用来说明一个页面或一篇文章主要内容的关键词或术语。Web 2.0 网站经常使用标签，将网站页面按不同的标签重新聚合。

在网站结构上，标签页面与网站分类页面类似。分类页面是列出本分类下的内容页面，标签页面是列出标有相同标签的内容页面。区别在于，标签并不存在分类那样的上下从属关系，如分类页面可以有一级产品分类，再分为二级产品分类，再向下还可以分

出更多。而标签不存在这种关系，每一个标签都是平等的，只不过有些标签下包含的页面可能更多。

补充材料（Supplemental Result）

补充材料是 Google 特有的一个名称，但相似现象可能存在于其他搜索引擎。所谓补充材料，是指 Google 把一些权重比较低的页面放在一个独立于主索引数据的补充索引库中。在补充索引库中的页面更新频率较低，排名到前面位置的机率也比较低。

2007 年以前，Google 在搜索结果中会标明哪些页面是补充材料。2007 年取消了这种标志，不过一般认为，补充材料本身还是存在的，只不过不再被标志而已。随着 Google 几次底层架构的更新换代，补充材料索引库的更新速度很可能有了很大提高。

垂直搜索（Vertical Search）

指的是专注于某个行业领域的搜索。垂直领域既可能是不同的主题，如生活搜索、购物搜索、交通搜索，也可能是不同的媒介形式，如视频、图片。

C 类 IP 地址（Class C IP Address）

一个 IP 地址如 198.197.196.195，其中第三个数字 196 被称为 C 类。

C 类数字相同的 IP 地址，通常是同一台服务器或处在同一网络上的服务器。所以如果两个网站 IP 地址第三个数字相同，如 198.197.196.195 和 198.197.196.194，搜索引擎会认为这两个网站之间是有一定关系的，很可能是在同一架服务器上。SEO 行业由此认为，具有同一个 C 类 IP 地址的网站或链接应该尽量避免，不然很容易被认为是站群或链接农场。

需要说明的是，这里所说的 C 类 IP 地址在技术上并不准确，只是在 SEO 行业的通行说法。

长尾关键词（Longtail Keywords）

单个搜索次数比较少，但总体数量巨大的非热门关键词。长尾这个术语源自于《连线》杂志主编 Chris Anderson 于 2004 年发表的系列文章，2006 年出版的《长尾》一书做了完整论述。

惩罚（Penalty）

网站因为使用不符合搜索引擎质量规范的手法，被搜索引擎给予排名下降处理。

CMS

内容管理系统，英文 Content Management System 的缩写。指的是用来创建和管理网站内容的软件。目前大部分网站都是使用 CMS 系统管理、数据库驱动的动态页面。

CPA

英文 Cost Per Action 的缩写，中文意译为每次行动成本。

CPA 也就是 PPA 模式中广告双方所商定的，每次用户完成一次行为的广告价格。

CPC

英文 Cost Per Click 的缩写，中文意译为每次点击成本。

CPC 就是 PPC 模式下，每次点击广告商所要支付的广告价格。

CPL

英文 Cost Per Lead 的缩写，中文意译为每次引导成本。

CPL 就是 PPL 模式中，用户完成一次引导，广告商所要付出的广告费用。

CPM

英文 Cost Per Mille 的缩写，中文意译为每千次显示成本，其中 M 是拉丁文一千次的缩写。

CPM 也就是 PPI 模式的广告价格。由于广告显示一次价格很低，所以通常按显示付费的广告是以每千次显示计价。

CPS

英文 Cost Per Sale 的缩写，中文意译为每次销售成本。

CPS 就是 PPS 模式下，用户完成一次购买后，广告商所要付出的广告价格。

导出链接（Outbound Links）

页面 A 上有一个链接指向页面 B，这个链接对页面 A 来说就是一个导出链接。

倒排索引（Inverted Index）

为了提高搜索引擎实时返回搜索结果的速度，直接用于排名的索引库不是正向索引，而是倒排索引。所谓倒排索引，是对索引库重新组织，形成一个从关键词到页面的映射。

地理定位（Geo-targeting）

搜索引擎根据用户 IP 地址判断出用户所在地理位置，返回更适合这个用户地理位置的搜索结果。有一些关键词搜索与地理位置有较强的关系，如天气、送餐、洗衣服务等。按地理定位提供相应的搜索结果，是搜索引擎改善用户体验的方法。SEO 人员利用这一

特性，有时候也需要对网站进行地理定位工作，使自己的网站在目标用户所在地理位置得到好的排名。

第一屏（Above The Fold）

直译是“第一次折叠之上”。指的是用户不需要拉动页面右侧滑动条或鼠标就能看到的网页最上部的内容。由于这是用户第一眼看到的网页内容，SEO 人员，或者更广泛地说，网络营销人员，应该把最重要的信息放在第一屏，在几秒钟内吸引住用户注意。

点击率（Click-through Rate）

用户实际点击一个搜索结果页面次数与这个搜索结果被展示总次数之比。在排名不变的情况下，提高点击率也就意味着提高流量。

动态 URL（Dynamic URL）

指的是包含有问号、等号及参数的 URL。如：

http://www.domain.com/index.php?catID=1&storyID=12345

通常动态 URL 对应的就是动态页面。问号、等号等字符后面所跟的参数就是所需要查询的数据库数据。

动态页面（Dynamic Pages）

与静态页面相对应。动态页面并不真实存在于服务器上，没有一个真正存在的文件对应。动态页面是由数据库驱动、程序脚本生成的页面。当用户访问动态页面时，程序查询数据库，并实时生成一个页面。目前大部分网站都是以动态页面为主。

Everflux

2003 年之后，Google 不再进行剧烈的 Google Dance，索引库数据以及排名算法更新都是不间断、小规模地随时启用，称为 Everflux。

反向链接（Back Links）

又称为导入链接（inbound links）。一个页面 A 上有一个链接指向页面 B，这个链接对页面 B 来说就是一个反向链接。

复制内容（Duplicate Content）

复制内容是指完全相同或非常相似的内容出现在多个页面上。复制内容既可能是由于转载、抄袭等原因出现在不同网站上，也可能是因为技术原因或网站结构方面的缺点而出现在同一个网站上。搜索引擎通常根据算法选出一个版本作为原创，其他页面上的相同内容被判断为复制内容，排名会受影响。

个人化搜索（Personalized Search）

也可翻译为个性化搜索。指的是搜索引擎根据用户个人信息返回不同的、更符合用户需求的搜索结果。引起不同搜索结果的因素包括浏览器设置、用户地理位置、用户搜索历史和网站浏览历史等。

工具条（Toolbar）

安装在浏览器上的插件，提供一些搜索引擎功能。用户可以在工具条上的搜索框内输入关键词直接搜索，而不必访问搜索引擎网站。几乎所有搜索引擎都开发了工具条供用户下载使用。

工具条 PR（Toolbar PR）

工具条 PR 指的是 Google 工具条显示的，站长可以查看的 Google PR 值。工具条 PR 值以数字 0～10 表示，0 最低，代表重要性最低，10 代表重要性最高的页面。要注意的是，工具条 PR 并不是 Google PR 值的绝对真实反映。

Google 保龄（Google Bowling）

恶意破坏竞争对手排名的一种方法。给竞争对手网站购买或群发大量垃圾链接，使 Google 误以为竞争对手作弊，从而惩罚竞争对手网站。虽然 Google 保龄发生机率很低，但现实中确实会发生。

Google Dance

2003 年以前，Google 每个月大规模更新索引库和排名算法，这个过程需要持续几天才能完成所有数据中心的更新。在 Google Dance 期间，很多网站的排名出现剧烈波动，访问不同数据中心看到的搜索结果也不同。目前 Google 已经不再 Dance 了。

Google 迷你全站链接（Google One-line Sitelinks）

顾名思义，这是 Google 全站链接的迷你版，不是显示 2 列 4 行 8 个链接，而是显示一行四个内页链接。

Google PR

Google PR 是 Google 通过链接关系计算出来的，用来衡量一个页面重要性的指标。其原理是把链接当做一个民主投票，页面 A 链接到页面 B，就意味着 A 对 B 进行了一次信任投票，使 B 的重要性提高。

决定 PR 值的既有反向链接的数量，也有链接质量。页面 A 本身的 PR 值高，传递给页面 B 的 PR 值也越高。另外，页面 A 上的导出链接总数也影响每个链接所能传递的 PR 值，页面 A 上导出链接越多，能分配和传递到页面 B 的 PR 值越低。

Google 全站链接（Google Sitelinks）

Google 给予权重比较高的网站的一种特殊排名显示格式。除了正常的页面标题、说明、URL 之外，还在结果下面按 2 列 4 行显示最多 8 个网站内页链接。

Google 炸弹（Google Bombing）

很多网站使用相同锚文字指向一个特定页面，虽然这个页面上并没有出现链接中的锚文字，但因为锚文字是 Google 排名算法的重要因素之一，被指向的页面在搜索这个锚文字时，还是能排到最前面，这种现象被称为 Google 炸弹。虽然称为 Google 炸弹，但其实主流搜索引擎都有这个现象。

关键词密度（Keyword Density）

页面上特定关键词出现次数与页面全部词数之比。

黑帽 SEO（Blackhat SEO）

黑帽 SEO 是指使用欺骗性的，违反搜索引擎规则的作弊手法，使网站排名提高。典型的黑帽 SEO 手法包括隐藏文字、隐藏链接、垃圾链接、桥页等。黑帽 SEO 违反搜索引擎质量规范，被搜索引擎发现时，通常会导致惩罚甚至网站完全被删除。

HTTP 头信息（Http Header）

用户浏览器或搜索引擎蜘蛛向服务器发出访问请求后，服务器所返回的响应消息中最前面定义信息特征的一段信息。比如下面是一段典型的头信息：

```
#1 Server Response: http://www.zaccode.com
HTTP Status Code: HTTP/1.1 200 OK
Date: Tue, 09 Feb 2010 06:36:19 GMT
Server: Apache/2.2.14 (Unix) mod_ssl/2.2.14 OpenSSL/0.9.8l DAV/2
mod_auth_passthrough/2.1 FrontPage/5.0.2.2635
X-Powered-By: PHP/5.2.11
Set-Cookie: uchome_auth=deleted; expires=Mon, 09-Feb-2009 06:36:18 GMT;
path=/
Vary: Accept-Encoding
Content-Length: 15596
Connection: close
Content-Type: text/html
```

其中定义了服务器状态码、页面返回时间、服务器类型、PHP 版本等信息。

HTTP 状态码（Http Status Code）

HTTP 状态码是 HTTP 头信息中最前面的一段，表明服务器响应的状态。如：

```
HTTP Status Code: HTTP/1.1 200 OK
```

常见的状态码如下。

200：表示一切正常，访问请求成功。

301：永久转向。

302：暂时转向。

404：文件不存在。

500：服务器内部错误。

检查 HTTP 状态码有助于 SEO 观察服务器是否工作正常，以及所设定的转向是否符合要求（是否是搜索引擎友好的 301 转向），页面不存在时是否返回 404 状态码等。

灰帽 SEO（Greyhat SEO）

灰帽 SEO 是介于黑帽 SEO 和白帽 SEO 之间，比较有争议性的 SEO 手法，比如链接买卖和站群等。这些手法既可能有益于用户，并遵守搜索引擎规则，也可能被滥用来进行欺骗性的排名。

降权

域名权重因为使用黑帽或灰帽手法而下降。降权既可能是搜索引擎算法自动甄别和处理，也可能是人工检查和处理。降权有点类似于惩罚，不同之处是，降权指的是整个域名排名能力下降，而惩罚有可能是针对特定页面或特定关键词。

交叉链接（Crosslink）

一组网站之间互相链接，目的是提高所有网站的外部链接数量。通常交叉链接的网站都是一个站长所控制，或同属于一个链接农场。这样得到的链接经常被搜索引擎认为是不自然的链接。

进入页面（Entry Page）

流量分析术语，指的是用户进入网站所访问的第一个页面。

静态 URL（Static URL）

不包含问号、#号、等号及参数的 URL，如：

http://www.domain.com/news/12345.html

原本静态 URL 是与静态页面相对应的。随着 SEO 观念的深入，以及 URL 重写技术的普遍应用，动态页面也可以实现静态 URL，称之为伪静态。

静态页面（Static Pages）

服务器上真实存在的文件对应的网站页面。无论哪个用户在什么时间访问这个页面，其内容都不会发生变化，除非站长在这个页面文件上的 HTML 代码中做了修改。

镜像网站（Mirrored Site）

一个或多个域名不同、但内容完全相同的网站。

镜像网站的出现有可能是有意的，也可能是无意的。有意的原因包括：

- 软件资源网站给用户提供多个下载链接。
- 某些网站提供多个镜像，使负载均衡或使服务更稳定。
- 黑帽 SEO 复制多个网站，意图获得更多排名。
- 整站被人复制抄袭也会造成镜像网站。

镜像网站通常会造成复制内容，搜索引擎并不喜欢，对原创内容网站或站长想排名的网站可能造成无法预知的影响。所以就算没有 SEO 意图的镜像网站，使用时也要小心。

绝对路径（Absolute Path）及相对路径（Relative Path）

页面需要链接到另一个网页时所使用的 HTTP 地址中，包含了域名的完整网址称为绝对路径。如果使用的是不包含域名的、被链接页面相对于当前页面的相对网址，则称为相对路径。

开放目录（DMOZ）

英文全称为 Open Directory Project，缩写为 ODP，意译即为开放目录项目。官方名称 DMOZ 源自于 Directory Mozilla，也就是目录中的 Mozilla。Mozilla 最初是网景浏览器（Netscape，最早的浏览器之一）的开发代号，现在已经演变为网上影响力最大的全球社区和非营利组织之一，也是很多开源软件的总称。

DMOZ 是一个只有几个管理人员，由招募的义务编辑来管理、审核网站的人工目录。开放目录的数据被包括 Google 等在内的很多网站使用。由于是人工编辑，在一定程度上保证了收录网站的质量，所以开放目录是少数几个很有价值的网站目录之一。

垃圾（Spam）

如名称所指明的，垃圾就是用户不需要的内容，如垃圾邮件。在搜索引擎及 SEO 行业中，垃圾指的是黑帽 SEO 纯粹为了获取排名创建的没有意义的内容及链接。

来路（Referral）

用户从页面 A 点击一个链接来到页面 B，则页面 A 就是页面 B 的一个来路，也就是说访问页面 B 的用户是点击页面 A 来的。

链接果汁（Link Juice）

链接果汁是一个比较笼统的概念，指的是链接所能传递到目标页面的权重。链接果汁多少取决于链接源页面本身权重、导出链接数目、页面是否被惩罚、是否被判定为付费链接等。

链接流行度（Link Popularity）

也称为链接广度、链接广泛度。指的是一个页面所获得的反向链接数量及质量的总和。链接流行度是页面重要性的指标之一。

链接农场（Link Farm）

没有任何实质内容，专门用来大量交换链接或给自己的网站制作链接的网站。链接农场网站页面上除了大量链接外就没有其他内容了。链接农场网站经常属于同一个站长或同一个网络。

链接农场网站及其上面的链接，目前都被搜索引擎认为是黑帽 SEO 手法。一旦被判断为链接农场，网站会被惩罚甚至删除。

链接 nofollow 属性

页面超链接的一个属性，也常简称为 NF 属性。也常俗称为 nofollow 标签。代码写法如下：

```
<a href="http://www.example.com/" rel="nofollow">这个链接将不被搜索引擎跟踪
</a>
```

链接加了 nofollow，就是告诉搜索引擎不要顺着这个链接爬行下去，链接权重也不传递。目前主流搜索引擎，包括百度、Google、雅虎、必应等都支持 NF 属性。

链接诱饵（Linkbait）

常用外部链接建设方法之一。创建吸引眼球，比如有争议、好玩、资源性、工具性的内容，吸引其他站长给予链接。

LSI

英文 Latent Semantic Indexing 的缩写，中文意译是潜在语义索引。

潜在语义索引指的是通过海量文献找出词汇之间的关系。当两个词或一组词大量出现在同一个文档中时，这些词之间就可以被认为是语义相关的。

锚文字（Anchor Text）

也就是链接文字，页面上超链接中可以点击的那段文字。以下面 http 代码为例：

```
<a href="http://www.zaccode.com/">这里是锚文字</a>
```

锚文字对于当前页面及被指向的页面主题都有很强的提示作用，对 SEO 有很大帮助。

Meta Nofollow 属性

页面 HTML 代码中 Meta 标签（元标签）的一种，格式如下：

```
<meta name="robots" content="nofollow" />
```

这个 Meta 指示搜索引擎不要跟踪和爬行这个页面上的所有链接。与链接 NF 属性不同的是，Meta 标签适用于于本页面上的所有链接，链接 NF 属性只对一个链接起作用。

MFA 网站

英文 Made For AdSense 缩写，也就是只为了做 Google AdSense 广告而存在的网站。Google 并不喜欢这种网站(至少 Google 表面上是这样说的)，虽然这些网站都参与 Google AdSense，也为 Google 赚了钱。

通常 MFA 网站没有什么实质内容，而是抄袭、采集其他网站文章，或用程序采集搜索引擎的搜索结果，自动生成大量页面，然后放上 Google AdSense 代码赚钱。绝大部分 MFA 网站用户体验很差。

面包屑导航（Breadcrumbs）

网站导航的一种，通常位于页面左上角，以一行文字链接的方式告诉用户目前所在的页面处于网站整体结构的哪个位置。页面面包屑导航包括了本页面的所有上级目录链接，所以用户可以一眼判断出自己当前所在位置。

免费链接网站（FFA）

Free For All 的缩写。六七年前挺流行的免费链接资源，任何人都可以在 FFA 网站上免费发布链接，最新发布的网站排在前面，旧的链接相应被推到越来越远的内页中。可以想象，有很多人在 FFA 网站不停的发布，希望能占据首页，所以大部分 FFA 首页几分钟就会被新发布的链接刷新。近几年来 FFA 的价值基本上为零，在 SEO 行业视野中已经消失了。

内部链接（Internal Links）

同一个域名之间的链接就是内部链接。与外部链接类似，内部链接也可以分为内部反向链接和内部导出链接。

内部优化（On-page Optimization）

或者称为页面上的优化，指的是在网站页面上进行完全由站长自己所控制的 SEO。如页面 Meta 标签的撰写和修改、网站结构和内部链接的优化等。

爬行和抓取（Crawl）

爬行指的是搜索引擎蜘蛛沿着超链接，从一个页面爬到另一个页面、发现更多网址的过程，就好像蜘蛛在蜘蛛网上爬行一样，蜘蛛及爬行都因此而得名。

搜索引擎蜘蛛发现新的 URL，就会把新 URL 上的内容记录下来存入数据库，称为抓取。

排名算法（Algorithm）

搜索引擎排名算法指的是用户输入关键词后，搜索引擎在自己的页面数据库中寻找、筛选，并且按一定规则对结果页面进行排名的过程。

PPA

英文 Pay Per Action 的缩写，中文意译为按行动付费。

PPA 是网络广告定价模式的一种，广告商在用户完成一个特定行动后，支付一定广告费用。这个特定行动可以是一次购买，或者是填写在线表格，订阅电子杂志，打电话联系广告商等。PPA 这种广告模式通常是使用在联署计划中。目前 Google 等搜索引擎也开始尝试 PPA 性质的广告。

PPC

PPC 是英文 Pay Per Click 的缩写，中文意译为按点击付费。

PPC 是一种网络广告模式。广告商每得到一次广告点击，就按商定的价格支付费用。虽然普通网站也可以按 PPC 模式卖广告，但网上使用 PPC 最为广泛的还是搜索竞价广告，包括搜索结果页面和内容发布网站上的搜索广告。通常在搜索引擎广告 PPC 模式中，广告商针对关键词进行竞价。

PPC 是 SEM 的重要组成部分之一。

PPI

英文 Pay Per Impression 的缩写，中文意译为按显示付费。

PPI 也是网络广告计价模式的一种，广告商的广告每显示一次，广告商就要付费，无论是否产生了点击、引导或销售。按显示付费是早期互联网广告的最重要形式，现在在主流门户网站上还占据很大份额。

PPL

英文 Pay Per Lead 的缩写，中文意译为按引导付费。

PPL 是 PPA（按行动付费）的一种，也就是用户完成一次引导行为，广告商就要付费。这里所说的引导通常是指没有购买，但与广告商发生联系的一次行为，诸如订阅电子杂志，注册为免费用户，填写在线联系表格，给广告商打一次电话等。这种引导行动比点击浏览网站更靠近完成销售。

PPS

英文 Pay Per Sale 的缩写，中文意译为按销售付费。

PPS 是 PPA 的一种，用户完成一次购买行为，广告商需要支付广告费用。

PR 劫持（PR Hijacking）

使用作弊手法将自己网站工具条 PR 值提高，通常是通过跳转（如 301 和 302 转向）实现。

QDF

英文 Query Deserves Freshness 的缩写，中文意思是“应该返回新鲜内容的搜索”。

Google 根据搜索趋势检测社会热点话题，与之相关的搜索词会更多返回新鲜的页面，也就是新创建的页面和刚刚更新的页面。

去重

搜索引擎分析一个网站的所有页面，消除存在于所有页面的重复部分，如导航、广告、版权声明等，提取页面上独特内容的过程。

Robots 文件

放在网站根目录下的一个纯文字文件 robots.txt，用来指示搜索引擎蜘蛛哪些页面可以被爬行和抓取。搜索引擎蜘蛛访问一个网站时，都要先打开 robots 文件，凡是 robots 文件指明禁止搜索引擎抓取的，搜索引擎就会忽略，不再爬行抓取。当然也有一部分恶意蜘蛛会忽略 robots 文件，因为它们的目的只是为了扫描邮件地址或抄袭文章，不会理睬 robots 文件。

三向链接

这是一种扩展的友情链接。比如网站 A 链接到网站 B，网站 B 链接到网站 C，网站 C 再链接回网站 A。

站长为了避免友情链接被搜索引擎检测出来而降低链接的效果，所以希望通过三向链接使外部链接看起来像是单向链接。其实搜索引擎很容易就可以检测到这种模式还是一种友情链接。

SEM

英文 Search Engine Marketing 的缩写，中文意译为搜索引擎营销。

SEM 是指在搜索引擎上推广网站，提高网站可见度，从而带来流量的网络营销活动。SEM 包括 SEO，PPC（搜索竞价排名），付费登录等形式，其中以 SEO 和 PPC 最为常见。

SEO

英文 Search Engine Optimization 的缩写，中文意译为搜索引擎优化。

SEO 是指在了解搜索引擎自然排名机制的基础上，对网站进行内部及外部的调整优化，改进网站在搜索引擎中的关键词自然排名，获得更多流量，从而达成网站销售及品牌建设的目标。

SERP

Search Engine Results Page 的缩写，意思是搜索引擎结果页面。用户输入关键词，点击搜索按钮后，搜索引擎返回显示的结果页面。

沙盒效应（Sandbox）

Google 对新网站的一种排名延迟处理方式。新网站在一段时间内无论如何优化，竞争度比较高的主要关键词都很难有好的排名，这段期间就称为沙盒。沙盒可以理解为 Google 给新网站的见习期。大部分主流搜索引擎都有类似效应，并不仅限于 Google。

删除（Ban）

网站因为严重作弊，所有页面被搜索引擎从数据库中删除，不予收录。

社会化媒体（Social Media）

指强调互动、参与以及用户产生内容的网站，比如博客、论坛、社交网络、内容分享、社会化书签网站等。社会化媒体网站没有了用户的参与和内容贡献，本身就没有了任何内容。

深度链接（Deep Links）

或者叫深层链接，指的是指向网站内页，而不是首页的外部链接。

枢纽网站（Hub）

枢纽网站指的是大量导出链接向高质量、高权威度的相关网站的网站。一般来说，枢纽网站本身也是权威度高的网站，因为这种网站内容主题集中，提供大量用户需要的资源链接。

SMM

英文 Social Media Marketing 的缩写，中文意译是社会媒体营销。

SMM 指的是在社会化媒体网站，如博客、线上社区、维基、视频分享网站、图片分享网站、社交网络等，进行营销和公关等活动。SMM 与 SEO 既有很大区别，也有互相促进和交叉的部分。

搜索引擎友好（Search Engine Friendly）

搜索引擎友好指的是搜索引擎容易爬行、抓取，容易提炼相关关键词的网站设计。要做到搜索引擎友好，涉及网站整体结构，内部链接，页面减肥，各种 HTML 代码的书写等。

索引（Index）

搜索引擎对抓取来的文件进行预处理，经过删除停止词、中文分词、关键词提取等过程，形成一个从页面到关键词集合的映射存入数据库，这个过程就叫索引，得到的数据库叫做索引库。

停止词（Stop Words）

在自然语言中出现频率非常高，但是对文章或页面的意义没有实质影响的那类词。如英文中的“the”，“and”，“of”等，中文中的“的”，“也”，“啊”等。停止词使用频繁，但对语义影响很小，搜索引擎遇到停止词时，无论是在索引或排名时，经常会忽略。忽略停止词对搜索排名几乎没有什么影响。

投资回报率（ROI）

Return On Investment 的缩写。ROI 指的是获得的收益与投入之比，这是衡量营销活动成功与否的最重要标志之一。在 PPC 营销中，ROI 的测量相对明确，因为每一个点击以及带来的销售数字都是有明确价值的。SEO 同样也有要达到的 ROI，只不过要计算的投入和收益，尤其是投入部分不是那么明确，比如投入的时间、人力成本，要转化为资金数字就不容易很准确，有很多因素要考虑。

图片 ALT 属性（Image ALT Text）

指的是网页上的图片因为某种原因不能被显示时应该出现的替代文字（alternative

text）。如下面这段代码所示：

```
<img src="images/pic.jpg" alt="这里就是 ALT 替代文字">
```

准确的说，图片替代文字是 ALT 属性，而不是 ALT 标签。但有时也常被称为 ALT 标签，属于大家约定俗成的称呼。

网站导航（Navigation）

页面上帮助用户明确目前所在位置，使用户能够比较容易地继续访问其他页面的一套链接系统。通常表现为页面顶部的菜单系统，左侧或右侧的导航条，页脚的辅助菜单等。一般情况下，网站导航系统与主要内容版块是一一对应的，有助于用户轻松找到相应内容。

转向（Redirect）

也称跳转、重定向。

转向是指当用户访问页面 A 时，被自动转移到页面 B，而用户并没有点击任何链接。转向可以由多种方式实现，如服务器端的 301 转向、302 转向，客户端的 JavaScript 转向，meta 刷新（meta refresh）等。跳转经常被黑帽 SEO 当做一种作弊手段。

URL 静态化

通过 URL 重写技术（URL rewrite），将动态 URL 转变为静态 URL。

在 LAMP（Linux+Apache+MySQL+PHP）主机上，URL 重写通常是通过 mod_rewrite 模块。在 Windows 主机上，通常是通过 ISAPI Rewrite 和 IIS Rewrite 模块。

外部链接（External Links）

不同域名之间的链接叫做外部链接。比如域名 A 上的任何一个页面 a，有链接指向域名 B 上的任何一个页面 b，这个链接对域名 A 和 B，以及页面 a 和 b 来说都是外部链接。

外部链接又可以分为外部反向链接和外部导出链接。上面所举的例子，对域名 A 来说是一个外部导出链接，对域名 B 来说，就是一个外部反向链接。对 SEO 人员来说，外部反向链接是影响排名的至关重要因素，所以外部链接也常常特指外部反向链接，也就是来自其他域名的反向链接。

外部优化（Off-page Optimization）

或者称为页面之外的优化。指的是不在网站页面本身上进行的 SEO，通常包括外部链接建设，社会化媒体网站的参与等。

网页快照（Cache）

搜索引擎数据库中记录的页面内容拷贝。搜索引擎在结果中给出“网页快照”（或其他类似称呼）链接，用户点击后看到的就是搜索引擎数据库中的页面内容。用户因为某种原因不能访问原始网页时，可以查看网页快照里的内容作为参考。

网站目录（Directory）

也称为网址站、地址站等。其他站长可以提交自己的网站，目录所有人审核批准或自己挑选收录网站。典型的网站目录如雅虎目录、开放目录、中文的好 123 等。被网站目录收录是建立外部链接的最常用手法之一。

网址规范化（URL Canonicalization）

同一个文件由于种种原因出现在同一个网站的不同 URL 上，搜索引擎需要判断哪一个 URL 是真正的、也就是规范化的网址。

信任指数（TrustRank）

信任指数源于斯坦福大学和雅虎的共同研究，是一个衡量网站垃圾程度的指标。其原理是，受信任的网站通常不会链接到垃圾网站，所以与信任指数高的网站点击距离近的网站，信任指数也高，距离越远，信任指数就越低。信任指数最高的是人工挑选出来的一组网站。

新闻源

新闻源是搜索引擎收集新闻的来源网站。被纳入新闻源的网站，不仅所发布的新闻会出现在搜索引擎新闻垂直搜索中，网站其他内容在普通页面搜索中也有比较高的权重。

要成为新闻源网站有一定的要求，而且要申请。百度新闻源申请在这里：

http://news.baidu.com/newsop.html

Google 在这里：

http://www.google.com/support/news_pub/bin/request.py?contact_type=suggest_content&hl=cn

XML 网站地图（XML Sitemap）

这是 Google 于 2005 年提出，并且得到大部分主流搜索引擎，如 Google、百度、雅虎、必应等支持的网站地图标准。所谓 XML 网站地图就是一个 XML 文件，在这个文件中列出网站上所有需要收录的 URL，还可以加上这些 URL 的信息，如更新日期、相对重要性等，通过这种方式通知搜索引擎，网站上有哪些 URL 需要收录。XML 网站地图对搜索引擎来说只是一种有益的参考，并不是收录的保证。

XML Sitemap 中的 Sitemap 这个词 S 必须大写，英文 SEO 文章中的 Sitemap 特指 XML 网站地图。相对应的，sitemap 通常是指网站上的 HTML 版本的网站地图。

页面劫持（Page Hijacking）

页面劫持是一种 SEO 作弊手法，指的是黑帽 SEO 使用各种手段，将本来应该访问页面 A 的用户，转向到另一个页面 B。

302 转向就是经常被用来实现页面劫持的手段。作弊者在自己的网站页面 A 上做 302 转向到其他网站上的页面 B。由于某种原因，搜索引擎把页面 A 当做原始页面，在搜索引擎排名中出现。用户点击结果 A 后，作弊者在 A 网站上通过程序把用户转向到完全无关的（通常是成人内容、赌博、卖各种违禁药品等的网站）页面。所以说页面劫持实际上是利用了其他网站上的高质量内容，用户却被劫持到自己的作弊网站上。

页面正文（Body Text）

SEO 领域中所说的正文并不是指 HTML 代码中<body></body>之间的内容，而是指排除页面导航、页脚、广告等之后的网页实质内容。

友情链接（Reciprocal Links，Exchanged Links）

又可以称为交换链接、互惠链接等。

友情链接是指 A，B 两个网站互相链接到对方，也就是说 A 网站链接到 B 网站，B 网站也链接到 A 网站。这是获得外部链接的最简单方式。

域名权重（Domain Authority）

一个域名在搜索引擎上排名的综合实力。域名权重是很多因素的总和，包括域名种类、历史、内容原创性、链接关系等。

着陆页（Landing Page）

进入页面的一种，不过着陆页着重于优化转化率的概念，而不是流量分析概念。着陆页指的是网络营销人员专门设计的吸引用户访问，并且通过各种手段提高用户转化的一组进入页面。

整合搜索（Universal Search）

也可翻译为通用搜索。指的是搜索引擎在搜索结果页面上同时显示多个垂直搜索内容，包括图片、视频、新闻、博客、财经等。整合搜索是所有主流搜索引擎近两年的趋势。

蜘蛛（Spider，Bot，Crawler）

也称为机器人。指的是搜索引擎运行的计算机程序，沿着页面上的超链接发现和爬行更多页面，抓取页面内容，送入搜索引擎数据库。

蜘蛛陷阱（Spider Trap）

蜘蛛陷阱指的是由于网站内部结构的某种特征，使搜索引擎蜘蛛陷入无限循环，无法停止爬行。最典型的蜘蛛陷阱是某些页面上的万年历，搜索引擎蜘蛛可以一直单击“下个月”陷入无限循环。

中文分词（Chinese Word Segmentation）

这是中文搜索特有的过程，指的是将中间没有空格的、连续的中文字符序列，分隔成一个一个单独的、有意义的单词的过程。在英文等拉丁文字中，词与词之间有空格自然区隔，所以没有分词的必要。而中文句子包含很多词，词之间没有自然分隔，搜索引擎在提取、索引关键词及用户输入了关键词需要进行排名时，都需要先进行分词。

转化率（Conversion Rate）

用户访问网站后，达成网站所定义的目标行动就称为一次转化，如完成订单、注册邮件列表、填写联系表格等。完成转化的用户数与所有访问用户数之比就称为转化率。

自然排名（Organic Ranking，Natural Ranking）

与付费和广告无关，只是依靠页面本身相关性、重要性而出现在搜索引擎结果页面的排名。在一个搜索引擎结果页面上，广告或付费排名通常都标有推广、赞助商链接之类的名称，自然排名则没有这些标记。

自愿链接（Editorial Links）

英文 editorial links 原意指的是有编辑意义的链接，也就是说其他站长因为你网站的内容有价值，而自愿链接到你的网站。editorial links 不容易直译，所以我把它称为自愿链接。自愿链接才是最有价值的链接，对网站排名帮助最多，不过获得自愿链接难度也比较大。

结束语：
下一步做什么

（1）如果你觉得本书对你稍有帮助，请在你的博客、网站上给“SEO 每天一贴”做个链接，网址是：

http://www.seozac.com//

我将继续更新博客，发布最新 SEO 动态。

（2）如果你已经有网站，从关键词研究开始，重新审视你的网站，一定能发现可以改进的地方。

如果你还没有网站，立即动手做一个网站，从关键词研究开始，到网站结构、页面优化、外链建设、流量分析，自己真正做一遍。第一个网站不一定成功，但不动手做就一定没有成功的时候。只要你认真读完本书，相信自己，你已经比大部分做网站的人更懂 SEO 了，尽快开始做网站吧。

（3）关注 SEO 最新消息和进展。下面是一些值得推荐的渠道。

中文 SEO 博客：

- 点石互动——http://www.dunsh.org/
- 天真的 SEM 博客——http://sem.la/blog/
- 樂思蜀 SEO 博客——http://www.lesishu.com/
- KYW 的 SEO 实验室——http://www.seotest.cn/
- 星箭 SEO 博客——http://www.starow.net/
- SEM Watch——http://semwatch.org/
- SEM 一家之言——http://www.semyj.com/
- Google 网管博客——http://googlewebmaster-cn.blogspot.com/

可惜的是，上面几个博客其实更新都很少很少了……

英文 SEO 博客：

- Search Engine Roundtable——http://www.seroundtable.com/
- SEOmoz——http://www.seomoz.org/blog
- Search Engine Land——http://searchengineland.com/
- Search Engine Journal——http://www.searchenginejournal.com/

- Google Webmaster Central——http://googlewebmastercentral.blogspot.com/
- Matt Cutts——http://www.mattcutts.com/blog
- Michael Gray——http://www.wolf-howl.com/
- SEObook——http://www.seobook.com/
- Bing Webmaster Blog——http://www.bing.com/community/blogs/webmaster/default.aspx

英文 SEO 博客比较多，上面是更新较多的几个。

中文 SEO 论坛：

- 点石论坛——http://www.dunsh.org/forums/index.php
- seobbs——http://www.seobbs.net/
- SEO 十万个为什么——http://www.seowhy.com/bbs/

SEO 行业会议：

- 年度点石大会及点石茶话会——http://www.dunsh.org/forums/forum-29-1.html
- 赢时代搜索营销大会——http://www.timev.com/event/
- MADCon——http://www.madcon.cn/

（4）本书读者交流、分享 SEO，请到专用网站 http://www.zaccode.com/

（5）需要 SEO 顾问咨询或培训服务，请参考 http://www.seozac.com/services/，或通过 E-mail 联系我：zanhui@gmail.com 或 me@zanhui.com

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

Broadview®
www.broadview.com.cn

《SEO 实战密码——60 天网站流量提高 20 倍（第 2 版）》读者交流区

尊敬的读者：

感谢您选择我们出版的图书，您的支持与信任是我们持续上升的动力。为了使您能通过本书更透彻地了解相关领域，更深入的学习相关技术，我们将特别为您提供一系列后续的服务，包括：

1. 提供本书的修订和升级内容、相关配套资料；
2. 本书作者的见面会信息或网络视频的沟通活动；
3. 相关领域的培训优惠等。

您可以任意选择以下四种方式之一与我们联系，我们都将记录和保存您的信息，并给您提供不定期的信息反馈。

1．在线提交

登录www.broadview.com.cn/　　，填写本书的读者调查表。

2．电子邮件

您可以发邮件至jsj@phei.com.cn或editor@broadview.com.cn。

3．读者电话

您可以直接拨打我们的读者服务电话：010-88254369。

4．信件

您可以写信至如下地址：北京万寿路173信箱博文视点，邮编：100036。

您还可以告诉我们更多有关您个人的情况，及您对本书的意见、评论等，内容可以包括：

（1）您的姓名、职业、您关注的领域、您的电话、E-mail地址或通信地址；

（2）您了解新书信息的途径、影响您购买图书的因素；

（3）您对本书的意见、您读过的同领域的图书、您还希望增加的图书、您希望参加的培训等。

如果您在后期想停止接收后续资讯，只需编写邮件“退订+需退订的邮箱地址”发送至邮箱：market@broadview.com.cn即可取消服务。

同时，我们非常欢迎您为本书撰写书评，将您的切身感受变成文字与广大书友共享。我们将挑选特别优秀的作品转载在我们的网站（www.broadview.com.cn）上，或推荐至CSDN.NET等专业网站上发表，被发表的书评的作者将获得价值50元的博文视点图书奖励。

更多信息，请关注博文视点官方微博：http://t.sina.com.cn/broadviewbj。

我们期待您的消息！

博文视点愿与所有爱书的人一起，共同学习，共同进步！

通信地址：北京万寿路 173 信箱　博文视点（100036）　　电话：010-51260888

E-mail：jsj@phei.com.cn，editor@broadview.com.cn

www.phei.com.cn
www.broadview.com.cn

反侵权盗版声明

电子工业出版社依法对本作品享有专有出版权。任何未经权利人书面许可，复制、销售或通过信息网络传播本作品的行为；歪曲、篡改、剽窃本作品的行为，均违反《中华人民共和国著作权法》，其行为人应承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。

为了维护市场秩序，保护权利人的合法权益，我社将依法查处和打击侵权盗版的单位和个人。欢迎社会各界人士积极举报侵权盗版行为，本社将奖励举报有功人员，并保证举报人的信息不被泄露。

举报电话：（010）88254396；（010）88258888
传　　真：（010）88254397
E-mail：　dbqq@phei.com.cn
通信地址：北京市万寿路 173 信箱
　　电子工业出版社总编办公室
邮　　编：100036